Unconventional kinetochore kinases KKT2 and KKT3 have unique centromere localization domains

Gabriele Marcianò†, Midori Ishii†, Olga O. Nerusheva†, and Bungo Akiyoshi*

Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK

†These authors contributed equally to this work

*For correspondence: bungo.akiyoshi@bioch.ox.ac.uk

Abstract

Chromosome segregation in eukaryotes is driven by the kinetochore, the macromolecular protein complex that assembles onto centromeric DNA and binds spindle microtubules. Cells must tightly control the number and position of kinetochores so that all chromosomes assemble a single kinetochore. A central player in this process is the centromere-specific histone H3 variant CENP-A, which localizes constitutively at centromeres and promotes kinetochore assembly. However, CENP-A is absent from several eukaryotic lineages including kinetoplastids, a group of evolutionarily divergent eukaryotes that have an unconventional set of kinetochore proteins. There are six proteins that localize constitutively at centromeres in the kinetoplastid parasite *Trypanosoma brucei,* among which two homologous protein kinases (KKT2 and KKT3) have limited similarity to polo-like kinases. In addition to the N-terminal kinase domain and the C-terminal divergent polo boxes, KKT2 and KKT3 have a central domain of unknown function as well as putative DNA-binding motifs. Here we show that KKT2 and KKT3 are important for the localization of several kinetochore proteins and that their central domains are sufficient for centromere localization in *T. brucei*. Crystal structures of the KKT2 central domain from two divergent kinetoplastids reveal a unique zinc-binding domain (termed the CL domain for centromere localization), which promotes its kinetochore localization in *T. brucei*. Mutations in the equivalent domain in KKT3 abolish its kinetochore localization and function. Our work shows that the unique central domains play a critical role in mediating the centromere localization of KKT2 and KKT3.

Introduction

The kinetochore is the macromolecular protein complex that drives chromosome segregation during mitosis and meiosis in eukaryotes. Its fundamental functions are to bind DNA and spindle microtubules (Musacchio and Desai, 2017). In most eukaryotes, kinetochores assemble within a single chromosomal region called the centromere. While components of spindle microtubules are highly conserved across eukaryotes (Wickstead and Gull, 2011; Findeisen et al., 2014), centromere DNA is known to evolve rapidly (Henikoff et al., 2001). Nevertheless, it is critical that a single kinetochore is assembled per chromosome and its position is maintained between successive cell divisions. A key player involved in this kinetochore specification process is the centromere-specific histone H3 variant CENP-A, which is found in most sequenced eukaryotic genomes (Talbert et al., 2009). CENP-A localizes specifically at centromeres throughout the cell cycle and recruits HJURP, a specific chaperone that incorporates CENP-A onto centromeres (Black and Cleveland, 2011; McKinley and Cheeseman, 2016; Stankovic and Jansen, 2017). Besides CENP-A, components of the constitutive centromere-associated network (CCAN) also localize at centromeres throughout the cell cycle. CENP-A-containing nucleosomes are recognized by CCAN components, which in turn recruit the KNL1-Mis12-Ndc80 (KMN) network that has microtubule-binding activities. In addition to these structural kinetochore proteins, several protein kinases are known to localize at mitotic kinetochores, including Cdk1, Aurora B, Bub1, Mps1, and Plk1 (Cheeseman and Desai, 2008). These protein kinases regulate various aspects of mitosis, including kinetochore assembly, error correction, and the spindle checkpoint (Carmena et al., 2012; London and Biggins, 2014; Hara and Fukagawa, 2018).

Kinetoplastids are evolutionarily divergent eukaryotes that are defined by the presence of a unique organelle called the kinetoplast that contains a cluster of mitochondrial DNA (d'Avila-Levy et al., 2015). Centromere positions have been mapped in three kinetoplastids: 20–120 kb regions that have AT-rich repetitive sequences in *Trypanosoma brucei* (Obado et al., 2007), ~16 kb GC-rich unique sequences in *Trypanosoma cruzi* (Obado et al., 2005), and ~4 kb regions in *Leishmania major* (Garcia-Silva et al., 2017). Although some DNA elements and motifs are enriched, there is no specific DNA sequence that is common to all centromeres in each organism, suggesting that kinetoplastids likely determine their kinetochore positions in a largely sequence-independent manner. However,

none of CENP-A or any other canonical structural kinetochore protein has been identified in kinetoplastids (Lowell and Cross, 2004; Berriman et al., 2005; Aslett et al., 2010). They instead have unique kinetochore proteins, such as KKT1–25 (Akiyoshi and Gull, 2014; Nerusheva and Akiyoshi, 2016; Nerusheva et al., 2019) and KKIP1–12 (D'Archivio and Wickstead, 2017; Brusini et al., 2019) in *T. brucei*. It remains unknown which of these proteins form the base of trypanosome kinetochores that recruits other proteins. There are six proteins that localize at centromeres throughout the cell cycle (KKT2, KKT3, KKT4, KKT20, KKT22, and KKT23), implying their close association with centromeric DNA. Indeed, we previously showed that KKT4 has DNA-binding activity in addition to microtubule-binding activity (Llauró et al., 2018; Ludzia et al., 2020). However, RNAi-mediated knockdown of KKT4 affected the localization of KKT20 but not other kinetochore proteins, suggesting that KKT4 is largely dispensable for kinetochore assembly.

In this study, we focused on KKT2 and KKT3, which are homologous to each other and have three domains conserved among kinetoplastids: a protein kinase domain that is classified as unique among known eukaryotic kinase subfamilies (Parsons et al., 2005), a central domain of unknown function, and divergent polo boxes (DPB). Presence of an N-terminal kinase domain and a C-terminal DPB suggests that KKT2 and KKT3 likely share common ancestry with polo-like kinases (Nerusheva and Akiyoshi, 2016). Interestingly, a protein kinase domain is not present in any constitutively-localized kinetochore protein in other eukaryotes, making these protein kinases a unique feature of kinetoplastid kinetochores. In addition to the three domains that are highly conserved among kinetoplastids, AT-hook and SPKK DNA-binding motifs are found in some species, suggesting that these proteins are located close to DNA (Akiyoshi and Gull, 2014). Although RNAi-mediated knockdown of KKT2 or KKT3 leads to growth defects (Akiyoshi and Gull, 2014; Jones et al., 2014), little is known about their molecular function. In this report, we have revealed a unique zinc-binding domain in the KKT2 and KKT3 central domain, which is important for their kinetochore localization and function in *T. brucei*.

Results

Localization of KKT2 and KKT3 is not affected by depletion of various kinetochore proteins

We previously showed in *T. brucei* procyclic form (insect stage) cells that kinetochore localization of KKT2 and KKT3 was not affected by KKT4 depletion (Llauró et al., 2018). To examine the effect of other kinetochore proteins for the recruitment of KKT2 and KKT3, we established RNAi-mediated knockdowns for KKT1, KKT6, KKT7, KKT8, KKT10/19, KKT14, KKT22, KKT23, KKT24, and KKIP1 (Figure 1A). Using inducible stem-loop RNAi constructs in cells expressing YFP-fusion of the target protein, we confirmed efficient depletion of YFP signals and observed severe growth defects for KKT1, KKT6, KKT8, KKT14, and KKT24. RNAi against KKT7, KKT10/19, and KKIP1 also caused severe growth defects as previously observed (D'Archivio and Wickstead, 2017; Ishii and Akiyoshi, 2020). In contrast, only a mild growth defect was observed for KKT23, while no growth defect was observed for KKT22. We next used these RNAi constructs in cells expressing either YFP-KKT2 or KKT3-YFP and found that these proteins formed kinetochore-like dots in all conditions (Figure 1B, C). These results show that KKT2 and KKT3 can localize at kinetochores even when various kinetochores proteins are depleted.

KKT2 and KKT3 are important for localization of some kinetochore proteins

We next examined whether KKT2 and KKT3 are important for the localization of other kinetochore proteins. KKT2 RNAi using a stem-loop construct caused growth defects, as previously reported (Figure 2A) (Ishii and Akiyoshi, 2020). We saw defective kinetochore localization for KKT14 upon induction of KKT2 RNAi, while other tested proteins still formed kinetochore-like dots at 1 day post-induction (Figure 2B, C). We next established RNAi against KKT3 (Figure 2D) and examined its effect on the localization of other kinetochore proteins at 2 days post-induction. Kinetochore localization of two constitutive kinetochore components, KKT22 and KKT23, was affected by KKT3 depletion (Figure 2E, F), while that of other tested proteins remained largely intact. It is noteworthy that KKT2 and KKT3 can apparently localize at kinetochores independently from each other. Together with the fact that KKT2 and KKT3 are homologous proteins, these results raised a possibility that KKT2/3 might have redundant roles in kinetochore assembly, so we next examined

the effect of double knockdown. Although the growth defect of the KKT2/3 double RNAi was not dramatically different from that of individual knockdowns (Figure 2G), which could be explained by residual KKT2/3 signals (Figure 2I), defective kinetochore localization was found for KKT1 and KKT4 at 1 day post-induction (Figure 2H, I). Taken together, these results show that KKT2 and KKT3 play important roles in recruiting multiple kinetochore proteins. It is possible that more efficient or rapid inactivation methods could reveal additional kinetochore proteins whose localizations depend on KKT2/3.

Multiple domains of KKT2 are able to localize at centromeres in *T. brucei*

To understand how KKT2 localizes at centromeres, we determined which domain was responsible for its centromere localization by expressing a series of truncated versions of KKT2, fused with a GFP-tagged nuclear localization signal peptide (GFP-NLS) in *T. brucei* (Figure 3A). We previously showed that ectopically-expressed KKT2 DPB (residues 1024–1260) localized at kinetochores (Nerusheva and Akiyoshi, 2016). The present study confirmed this result and also identified two other regions (residues 562–677 and 672–1030) that localized at kinetochores from S phase to anaphase (Figure 3A, B and Figure S1).

Based on our previous finding that KKT2 co-immunoprecipitated with a number of kinetochore proteins (Akiyoshi and Gull, 2014), we reasoned that these fragments might localize at kinetochores by interacting with other kinetochore proteins. To test this possibility, we immunoprecipitated KKT2 fragments and performed mass spectrometry to identify co-purifying proteins. Although the central domain co-purified only with limited amounts of KKT3 and KKT5 (Figure 3C and Table S1), KKT2$^{672–1030}$ co-purified with several kinetochore proteins, with KKIP1 being the top hit (Figure 3D and Table S1). Similarly, KKT2 DPB co-purified with several kinetochore proteins, which was abolished in the W1048A mutant that did not localize at kinetochores (Figure 3E and Table S1) (Nerusheva and Akiyoshi, 2016). These results support the possibility that ectopically-expressed KKT2$^{672–1030}$ and DPB are able to localize at kinetochores from S phase to anaphase by interacting with non-constitutive kinetochore proteins (e.g. KKT1, KKT6, KKT7, KKT8,

KKIP1). A corollary is that, in wild-type cells, the constitutively-localized KKT2 protein recruits these transient kinetochore proteins onto kinetochores using KKT2$^{672–1030}$ and DPB domains.

The central domain of KKT3 is able to localize at centromeres constitutively

We next expressed KKT3 fragments in trypanosomes (Figure 4A, B). Similarly to KKT2, the N-terminal protein kinase domain of KKT3 did not localize at centromeres. KKT3 DPB had robust kinetochore localization only during anaphase (Figure 4C), which differs from KKT2 DPB that localized from S phase to anaphase. Immunoprecipitation of KKT3 DPB identified a number of co-purifying kinetochore proteins (Figure 4D), raising a possibility that KKT3, like KKT2, recruits other kinetochore proteins by its DPB.

KKT3$^{594–1058}$ and KKT3$^{594–811}$ that contain the central region also localized at kinetochores. Kinetochore localization was also observed for KKT3$^{594–728}$, which lacks AT-hook and SPKK DNA-binding motifs (residues 771–780). KKT3$^{594–811}$ co-purified with limited amounts of KKT7, KKT1, and KKT2 (Figure 4E). Importantly, KKT3$^{594–728}$ constitutively localized at kinetochores (Figure S1), suggesting that the central domain plays a crucial role in recruiting KKT3 onto centromeres throughout the cell cycle.

The *Bodo saltans* KKT2 central domain adopts a unique structure

To gain insights into how the central domains of KKT2 and KKT3 localize at centromeres, we expressed and purified recombinant proteins for their structure determination by X-ray crystallography. Our attempts to purify the *T. brucei* KKT3 (*Tb*KKT3 hereafter) central domain were unsuccessful, but we managed to express and purify from *E. coli* the central domain of KKT2 from several kinetoplastids, including *Bodo saltans* (a free-living kinetoplastid (Jackson et al., 2016)) and *Perkinsela* (endosymbiotic kinetoplastids (Tanifuji et al., 2017)) (Figure S2). We obtained crystals of *Bs*KKT2$^{572–668}$ (which corresponds to residues 569–664 in *T. brucei* KKT2) and determined its structure to 1.8 Å resolution by zinc single-wavelength anomalous dispersion (Zn-SAD) phasing (Figure 5 and Table 1). Our analysis revealed the presence of two distinct zinc-binding domains: the N-terminal one (referred to as the "CL" domain for its key role in centromere localization: see below)

consists of 2 β-sheets (where β-strands 1, 4 and 5 comprise the first β-sheet, and β-strands 2 and 3 comprise the second β-sheet) and one α-helix, while the C-terminal one consists of one β-sheet (comprising β-strands 6 and 7) and one α-helix (Figure 5A, B). The CL domain coordinates two zinc ions and the C-terminal domain coordinates one zinc ion.

A structural homology search using the DALI server (Holm and Laakso, 2016) indicated that the CL domain has weak structural similarity to proteins that have C1 domains (Table S2). C1 domains were originally discovered as lipid-binding modules in protein kinase Cs (PKCs) and are characterized by the $HX_{12}CX_2CXnCX_2CX_4HX_2CX_7C$ motif (Colón-González and Kazanietz, 2006; Das and Rahman, 2014). C1 domains are classified into a typical C1 domain, which binds diacylglycerol or phorbol esters, and an atypical C1 domain, which is not known to bind ligands. The closest structural homolog of the *Bs*KKT2 CL domain was the atypical C1 domain of the Vav1 protein (RMSD 2.7 Å across 52 Cα). Although the CL domain and the C1 domain share some structural similarity, their superposition revealed fundamental differences (Figure 6). Coordination of one zinc ion in the *Bs*KKT2 CL domain occurs via the N-terminal residues Cys 580 and His 584, while that in the Vav1 C1 domain occurs via the N-terminal His 516 and C-terminal Cys 564. More importantly, the CL domain does not have the $HX_{12}CX_2CXnCX_2CX_4HX_2CX_7C$ motif that is present in all C1 domains. Therefore, the structural similarity of these two distinct domains is likely a product of convergent evolution.

Structural analysis of the C-terminal zinc-binding domain of *Bs*KKT2 revealed a classical C2H2-type zinc finger (Table S3). C2H2 zinc fingers are known to bind DNA, RNA, or protein (Krishna et al., 2003; Brayer and Segal, 2008). In most known cases, two or more C2H2 zinc fingers are used to recognize specific DNA sequences, which is typically achieved by specific interactions between the side chain of residues in positions -1, 2, 3, and 6 in the recognition α-helix (where -1 is the residue immediately preceding the α-helix) and DNA bases (Wolfe et al., 2000). Notably, some proteins with a single zinc finger can recognize specific DNA sequences (Omichinski et al., 1997; Dathan et al., 2002). The C-terminal zinc-binding domain of *Bs*KKT2 consists of one C2H2 domain (-1: Ser 653, 2: Thr 655, 3: Lys 656, 6: Tyr 659). The sequence alignment of the *Bs*KKT2 C2H2 zinc

finger shows that residues at the position -1 and 3 are highly conserved in trypanosomatids, while those at position 2 and 6 are not (Figure 8A).

The CL domain structure is conserved in *Perkinsela* KKT2a

We next asked whether the central domain structure is conserved among kinetoplastids. *Perkinsela* is a highly divergent endosymbiotic kinetoplastid that lives inside *Paramoeba* (Tanifuji et al., 2017). Our homology search identified three proteins that have similarity to KKT2 and KKT3. Interestingly, similarities among these *Perkinsela* proteins are higher than those between them and KKT2 or KKT3 in other kinetoplastids (Figure S3). Because these *Perkinsela* proteins overall have higher sequence similarity to KKT2 than KKT3, we call them *Pk*KKT2a (XU18_4017), *Pk*KKT2b (XU18_0308), and *Pk*KKT2c (XU18_4564). *Pk*KKT2c does not have a kinase domain, like KKT20 in other kinetoplastids (Nerusheva and Akiyoshi, 2016). Our sequence alignment suggests that *Pk*KKT2a and *Pk*KKT2b have a CL-like domain but lack a C2H2 zinc finger (Figure S3).

We determined the crystal structure of *Pk*KKT2a$^{551–679}$ at 2.9 Å resolution by Zn-SAD phasing (Figure S2 and Table 2), which confirmed the presence of a CL-like structure: 2 β-sheets (residues 551–647), followed by an extended C-terminal α-helix (residues 648–679) (Figure 7A, B). The CL-like domain of *Pk*KKT2a overlaps closely with that of *Bs*KKT2 (RMSD: 0.79 Å across 39 Cα), with the exception of some differences being localized to the loop insertion and the absence of a C2H2 domain in *Pk*KKT2a$^{551–679}$ (Figure 7C), consistent with our sequence analysis (Figure S3). Taken together, our structures have revealed that the CL domain is conserved in *Bs*KKT2 and *Pk*KKT2a. Given the sequence similarity of KKT2 between *Bodo saltans* and other trypanosomatids (Figure 8A), it is likely that the unique CL domain structure is conserved among kinetoplastids.

***Perkinsela* KKT2a has DNA-binding activity**

Although *Perkinsela* KKT2a lacks a C2H2 zinc finger, our sequence analysis of the KKT2 central domain revealed a putative C2H2 zinc finger not only in trypanosomatids and bodonids but also in one of Prokinetoplastina's KKT2-like proteins, PhM_4_m.86555 (Figure S3). Furthermore, DNA-

binding SPKK motifs (Suzuki, 1989) are present right after the C2H2 zinc finger in many kinetoplastids, while an AT-hook motif is present within the CL-like domain of *Perkinsela* KKT2b (Figure S3). These observations suggest that the central domain of KKT2 and KKT3 have DNA-binding activity, perhaps stabilizing their localization at centromeres. To test this hypothesis, we performed fluorescence polarization assays using fluorescently-labelled DNA probes. Unfortunately, we were unable to obtain reliable data for *T. brucei* and *B. saltans* KKT2 central domains due to fluorophore quenching. We therefore focused on *Pk*KKT2a$^{551–679}$ that has a CL-like domain. Our fluorescence polarization assay showed that *Pk*KKT2a$^{551–679}$ has DNA-binding activity with a K_d of ~500 nM on three DNA probes that have different GC contents (50 mer CEN is 50-bp DNA sequence from the CIR147 centromere repeat in *T. brucei*) (Figure 7D). To assess the importance of the structural integrity for DNA binding, we next performed fluorescence polarization assay for *Pk*KKT2a$^{551–679}$ that has mutations in Zn-coordinating residues (C646 and C649, which correspond to C616 and C619 in *Tb*KKT2) and found that it has a similar DNA-binding affinity compared to wild-type *Pk*KKT2a$^{551–679}$ (Figure 7E). As a comparison, we used a well-characterized zinc finger (Designed zinc finger) that binds a specific DNA sequence (Designed DNA) (Jantz and Berg, 2010). This protein bound its optimal DNA sequence with a K_d of 8 nM, while it had weaker affinity for 20-bp and 50-bp probes from the CIR147 centromere sequence (Figure 7F). These results show that *Pk*KKT2a$^{551–679}$ has DNA-binding activity. It is possible that *Pk*KKT2a$^{551–679}$ has higher DNA-binding activity for *Perkinsela* centromere DNA sequences (yet to be identified).

The CL domain of KKT2 is important for long-term viability in *T. brucei*

To examine the functional relevance of the CL domain and C2H2 zinc finger, we tested their mutants in *T. brucei*. We first made various mutants in full-length *Tb*KKT2 and found that all mutants localized at kinetochores (Figure S4). Because *Tb*KKT2 has multiple domains that can independently localize at kinetochores (Figure 3), we next expressed mutants in our ectopic expression of the central domain (*Tb*KKT2$^{562–677}$). We found that mutations in Zn-coordinating residues of the CL domain (C576A, H580A, C597A, C600A, C616A, C619A) all abolished kinetochore localization (Figure 8A,

B). In contrast, similar mutations in the C2H2-type zinc finger (C640A C643A) did not affect the localization.

To gain insights into how the KKT2 CL domain may promote kinetochore localization, we analyzed conservation and electrostatic potential of the *Bs*KKT2 CL domain surface residues to identify possible patches that may be involved in this process. Our analysis revealed a highly conserved acidic patch, centered around residue *Bs*KKT2 D626 (Figure 5C, D, and Figure S5). Interestingly, this aspartic acid is strictly conserved in all KKT2 and KKT3 proteins (Figure S3). To test the importance of this residue, we mutated the corresponding residue in *T. brucei* and found that *Tb*KKT2$^{562–677}$ with either D622A or D622E failed to localize at kinetochores (Figure 8B). Taken together, our results show that the CL domain, but not the C2H2 zinc finger, is important for the kinetochore localization of the *Tb*KKT2 central domain.

To test the importance of the *Tb*KKT2 central domain for cell viability, we next performed rescue experiments. We replaced one allele of *Tb*KKT2 with an N-terminally YFP-tagged *Tb*KKT2 construct that has either wild-type or mutant versions of the central domain, and performed RNAi against the 5′UTR of the *Tb*KKT2 transcript to knockdown the untagged allele of *Tb*KKT2 (Figure S6) (Ishii and Akiyoshi, 2020). As expected, mutants in the CL domain (C576A and D622A) and the C2H2 zinc finger (C640A C643A) both localized at kinetochores (Figure 8C and Figure S7). Upon induction of RNAi, however, the CL domain mutants failed to support normal cell growth after day 4, while the C2H2 zinc finger mutants rescued the growth defects. These data confirm the importance of the CL domain for the function of *Tb*KKT2 *in vivo*.

Localization of KKT3 depends on the central domain in *T. brucei*

The central domain of *Tb*KKT3 can localize at kinetochores throughout the cell cycle (Figure 4). The sequence similarity of the central domain between KKT2 and KKT3 suggested that *Tb*KKT3 likely consists of two domains, which correspond to the CL domain and the C2H2 zinc finger present in KKT2 (Figure S3). Consistent with this prediction, mutating *Tb*KKT3 residues that align with Zn-coordinating histidine or cysteine residues in the CL domain of KKT2 abolished the kinetochore localization of the ectopically-expressed full length *Tb*KKT3 protein (Figure 9A, B). We also found

that the conserved aspartic acid D692 was essential for kinetochore localization. In contrast, mutations in the *Tb*KKT3 C2H2 zinc finger (C707A C710A) did not affect the kinetochore localization.

We next performed rescue experiments by replacing one allele of *Tb*KKT3 with a C-terminally YFP-tagged construct that has either wild-type or mutant versions of the central domain, and performed RNAi against the 3′UTR of *Tb*KKT3 to knockdown the untagged allele of *Tb*KKT3 (Figure S6). We first confirmed that *Tb*KKT3 CL domain mutants (C668A C671A and D692A) were unable to localize at kinetochores, while the *Tb*KKT3 C2H2 zinc finger mutant (C707A C710A) localized normally (Figure 9C and Figure S7). Upon induction of RNAi, CL mutants failed to rescue the growth defect, showing that kinetochore localization is essential for the *Tb*KKT3 function (Figure 9C). In contrast, the *Tb*KKT3 C2H2 zinc finger mutant supported normal cell growth. These data show that the *Tb*KKT3 CL domain is essential for the localization and function of *Tb*KKT3.

Discussion

A major open question concerning the biology of kinetoplastids is how these organisms assemble kinetochores specifically at centromeres using a unique set of kinetochore proteins. Studies in other eukaryotes have shown that constitutively localized kinetochore proteins (such as CENP-A and CENP-C) play crucial roles in kinetochore specification and assembly (French and Straight, 2017; Hamilton and Davis, 2020; Kixmoeller et al., 2020). Among the six proteins that constitutively localize at kinetochores in *T. brucei* (KKT2, KKT3, KKT4, KKT20, KKT22, and KKT23), we previously showed that KKT4 is important for recruiting KKT20 but not many other proteins including KKT2 and KKT3 (Llauró et al., 2018). In this study, we show that KKT2 and KKT3 are important for recruiting multiple kinetochore proteins (including KKT4, KKT22, and KKT23), while their localization is independent from various kinetochore proteins. Together with the fact that KKT2 and KKT3 have DNA-binding motifs, these results support a hypothesis that KKT2 and KKT3 locate at the base of kinetoplastid kinetochores and play crucial roles in recruiting other kinetochore proteins. It will be important to identify which proteins directly interact with KKT2/3 to understand the mechanism of kinetochore assembly.

KKT2/3 have sequence similarity with polo-like kinases (Nerusheva and Akiyoshi, 2016). In addition to an N-terminal protein kinase domain and a C-terminal divergent polo boxes, they have a central domain that is highly conserved among kinetoplastids. By ectopically expressing fragments of KKT2 and KKT3 in *T. brucei*, we established that their central domains can localize specifically at centromeres. The crystal structure of the *Bodo saltans* KKT2 central domain revealed a unique structure, which consists of two distinct zinc-binding domains, the CL domain and C2H2-type zinc finger. It is likely that the central domain of *T. brucei* KKT2 has a similar structure based on high sequence similarity between *Bs*KKT2 and *Tb*KKT2 proteins. Importantly, mutational analyses of *Tb*KKT2 revealed that the CL domain is important for the localization of the central domain, while the C2H2 zinc finger is not. Furthermore, although full-length *Tb*KKT2 CL mutants localized at kinetochores (likely due to interactions with other kinetochore proteins via other domains of *Tb*KKT2), they were not fully functional. Taken together, these data have established that the CL domain is essential for the function of *Tb*KKT2, which is consistent with the presence of CL, but not C2H2 zinc finger, in *Perkinsela* KKT2a. It remains unclear whether the structure of the central domain is conserved between KKT2 and KKT3. Nonetheless, our functional studies showed that the equivalent domain of CL in *Tb*KKT3 was also essential for the kinetochore localization and function, while the equivalent domain of C2H2 zinc finger was not, showing that the functional importance of CL is conserved in KKT3. It will be important to obtain KKT3 central domain structures to reveal structural similarity or difference between KKT2 and KKT3. It is noteworthy that all identified KKT2/3 homologs in deep-branching Prokinetoplastina have higher similarity to KKT2 than KKT3. We speculate that ancestral kinetoplastids had only KKT2-like proteins that carried out all necessary functions and that KKT3 in trypanosomatids and bodonids represents a product of gene duplication that became specialized in certain functions (such as more efficient centromere localization by its central domain compared to KKT2).

It remains unclear how kinetoplastids specify kinetochore positions. The fact that KKT2/3 central domains manage to localize at centromeres suggests that they are able to recognize something special at centromeres. What might be a unique feature at centromeres in kinetoplastids that lack CENP-A? Histone variants are one possibility. *T. brucei* has four histone variants, H2AZ, H2BV,

H3V, and H4V. However, none of them is specifically enriched at centromeres (Lowell and Cross, 2004; Lowell et al., 2005; Siegel et al., 2009), and histone chaperones did not co-purify with any kinetochore protein (Akiyoshi and Gull, 2014). Alternatively, there might exist certain post-translational modifications on histones or DNA specifically at centromeres (e.g. phosphorylation, methylation, acetylation, ubiquitination, or sumoylation). The KKT2 CL domain has a highly conserved acidic patch, which might act as a "reader" for such modifications. Although there is no known histone or DNA modification that occurs specifically at centromeres, KKT2/3 have a protein kinase domain and KKT23 has a Gcn5-related N-acetyltransferase (GNAT) domain (Nerusheva et al., 2019). It will be important to examine whether these enzymatic domains are important for proper recruitment of KKT2/3 central domains. Another unique feature at centromeres is the presence of kinetochore proteins, which could potentially recruit newly synthesized kinetochore components by direct protein-protein interactions. Finally, it is important to note that it remains unclear whether kinetoplastid kinetochores build upon nucleosomes. It is formally possible that the KKT2/3 central domains directly bind DNA and form a unique environment at centromeres. Understanding how the KKT2/3 central domains localize specifically at centromeres will be key to elucidating the mechanism of how kinetoplastids specify kinetochore positions in the absence of CENP-A.

Materials and Methods

Tryp cells and plasmids, microscopy, immunoprecipitation, and mass spectrometry

All trypanosome cell lines, plasmids, primers, and synthetic DNA used in this study are listed in Table S4. All trypanosome cell lines used in this study were derived from *T. brucei* SmOxP927 procyclic form cells (TREU 927/4 expressing T7 RNA polymerase and the tetracycline repressor to allow inducible expression) (Poon et al., 2012). Cells were grown at 28˚C in SDM-79 medium supplemented with 10% (v/v) heat-inactivated fetal calf serum (Brun and Schönenberger, 1979). Endogenous YFP tagging was performed using the pEnT5-Y vector (Kelly et al., 2007). Endogenous tdTomato tagging was performed using pBA148 (Akiyoshi and Gull, 2014) and its derivatives.

Inducible expression of GFP-NLS fusion proteins was performed using pBA310 (Nerusheva and Akiyoshi, 2016).

To make pBA1711 (KKT3 3′UTR hairpin RNAi construct targeting the KKT3 transcript from stop codon to +370 bp), BAG95 synthetic DNA fragment was digested with HindIII/BamHI and subcloned into the HindIII/BamHI sites of pBA310. pBA2052 (KKT2/3 double hairpin RNAi construct targeting KKT2 5′UTR from -342 bp to start codon as well as KKT3 3′UTR from stop codon to +370 bp) was made with BAG129 as above. Similarly, pBA861 (KKT1 hairpin RNAi targeting 1351–1776 bp) was made with BAG24, pBA864 (KKT6 hairpin RNAi targeting 12–499 bp) with BAG27, pBA869 (KKT14 hairpin RNAi targeting 859–1274 bp) with BAG32, pBA1316 (KKT8 3′UTR hairpin RNAi targeting from +31 to +446 bp) with BAG79, pBA1845 (KKT22 hairpin RNAi targeting 466 bp to stop codon) with BAG105, pBA1997 (KKT24 hairpin RNAi targeting 1001–1400 bp) with BAG115, and pBA2021 (KKT23 hairpin RNAi targeting 635–1047 bp) with BAG127. To make pBA1091 (KKIP1 RNAi), 297–856 bp of KKIP1 coding sequence was amplified with primers BA1541/BA1543 and cloned into p2T7-177 using BamHI/HindIII sites (Wickstead et al., 2002). To make pBA1807 (C-terminal YFP-tagging of KKT3), 4–3174 bp of KKT3 coding sequence and 250 bp of 3′UTR were amplified with primers BA2351/BA2352 and BA2353/BA2354, digested with HindIII/NotI and NotI/SpeI, respectively, and were cloned into the pEnT5-Y using HindIII/SpeI sites. Details of other plasmids are described in Table S4. Site-directed mutagenesis was performed using primers and template plasmids listed in Table S4. All constructs were sequence verified.

Plasmids linearized by NotI were transfected into trypanosomes by electroporation into an endogenous locus (pEnT5-Y derivatives and pBA148/pBA192/pBA892 derivatives) or 177 bp repeats on minichromosomes (pBA310 derivatives and p2T7-177 derivatives). Concentrations of drugs used were as follows: 5 µg/ml for phleomycin, 25 µg/ml for hygromycin, 10µg/ml for blasticidin, 30 µg/ml for G418, and 1 µg/ml for puromycin. To obtain endogenously-tagged clonal strains, transfected cells were selected by the addition of appropriate drugs and cloned by dispensing dilutions into 96-well plates. Clones that express mutant versions of KKT2 or KKT3 from the endogenous locus were screened by Sanger sequencing of genomic DNA. Expression of GFP-NLS fusion proteins (pBA310

derivatives) was induced by the addition of doxycycline (10 ng/ml). RNAi was induced by the addition of doxycycline (1 µg/ml).

Fluorescence microscopy, immunoprecipitation of GFP-fusion proteins, and mass spectrometry were performed essentially as described previously (Nerusheva and Akiyoshi, 2016; Ishii and Akiyoshi, 2020). Proteins identified with at least two peptides were considered as significant and shown in Table S1.

Multiple sequence alignment

Protein sequences and accession numbers for KKT2 and KKT3 homologs were retrieved from TriTryp database (Aslett et al., 2010), Wellcome Sanger Institute (https://www.sanger.ac.uk/), UniProt (UniProt Consortium, 2019), or a published study (Butenko et al., 2020). Searches for KKT2/3 homologs in Prokinetoplastina and Bodonida were done using hmmsearch on its predicted proteome using manually prepared KKT2/3 hmm profiles (HMMER version 3.0 (Eddy, 1998)). Multiple sequence alignment was performed with MAFFT (L-INS-i method, version 7) (Katoh et al., 2019) and visualized with the Clustalx coloring scheme in Jalview (version 2.10) (Waterhouse et al., 2009).

Protein expression and purification

Multiple sequence alignment together with secondary structure predictions of the KKT2 central domain were used to design constructs in *Bodo saltans* and *Perkinsela*. To make pBA1660 (*Bs*KKT2$^{572–668}$ with an N-terminal TEV-cleavable hexahistidine (His_6) tag), the central domain of *Bodo saltans* KKT2 (accession number BSAL_50690) was amplified from BAG50 (a synthetic DNA that encodes *Bodo saltans* KKT2, codon optimized for expression in Sf9 insect cells (Table S4)) with primers BA2117/BA2118 and cloned into the RSFDuet-1 vector (Novagen) using BamHI/EcoRI sites with an NEBuilder HiFi DNA Assembly Cloning Kit (NEB) according to the manufacturer's instructions. To make pBA1139 (His_6-*Pk*KKT2$^{551–679}$), the central domain of *Perkinsela* CCAP 1560/4 KKT2a (accession number XU18_4017) was amplified from BAG48 (a synthetic DNA that encodes *Perkinsela* KKT2a, codon optimized for expression in Sf9 insect cells (Table S4)) with

primers BA1569/BA1570 and cloned into RSFDuet-1 using BamHI/EcoRI sites with an In-Fusion HD Cloning Plus kit (Takara Bio). To make pBA2276 (His_6-Designed zinc finger), Designed zinc finger domain was amplified from BAG136 (a synthetic DNA codon optimized for expression in *E. coli*) with primers BA3077/BA3078 and cloned into RSFDuet-1 using BamHI/EcoRI sites with an NEBuilder HiFi DNA Assembly Cloning Kit. Recombinant proteins were expressed in BL21(DE3) *E. coli* cells at 20˚C using auto induction media (Formedium) (Studier, 2005).

Briefly, 500 mL of cells were grown at 37˚C in 2.5 L flasks at 300 rpm until OD_{600} of 0.2–0.3 and then cooled down to 20˚C overnight (2 L for *Bs*KKT2, 6 L for *Pk*KKT2a, and 2L for Designed zinc finger). Cells were harvested by centrifugation and resuspended in 50 ml per litre of culture of lysis buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM TCEP, 10 mM imidazole, and 1.2 mM PMSF). Proteins were extracted by mechanical cell disruption using a French press (1 passage at 20,000 PSI) and the resulting lysate was centrifuged at 48,384 g for 30 min at 4˚C. Clarified lysate was incubated with 5 mL TALON beads (Takara Bio), washed with 150 mL lysis buffer and eluted in 22 mL of elution buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM TCEP, and 250 mM imidazole) in a gravity column, followed by TEV treatment for the removal of the His_6 tag. Salt concentration of the sample was subsequently reduced to 50 mM NaCl using buffer A (50 mM Hepes pH 7.5, and 1 mM TCEP) and the sample was loaded onto a 5 mL HiTrap Heparin HP affinity column (GE healthcare) pre-equilibrated with 5% buffer B (50 mM Hepes pH 7.5, 1 M NaCl, and 1 mM TCEP) on an ÄKTA pure 25 system. Protein was eluted by using a gradient from 0.05 to 1 M NaCl, and protein-containing fractions were combined, concentrated with an Amicon stirred cell using an ultrafiltration disc with 10 kDa cut-off (Merck), and then loaded onto a HiPrep Superdex 75 16/60 size exclusion chromatography column (GE healthcare) pre-equilibrated with 25 mM Hepes pH 7.5, 150 mM NaCl, and 1 mM TCEP. Fractions containing the protein of interest were pooled together, concentrated with an Amicon stirred cell using an ultrafiltration disc with 10 kDa cut-off, and stored at -80˚C. Designed zinc finger was buffered exchanged into 50 mM Tris pH 7.5, 1 mM $ZnCl_2$, 50 mM NaCl and 1 mM TCEP prior storage. Protein concentration was measured by Bradford assay.

Crystallization

Both *Bs*KKT2$^{572–668}$ and *Pk*KKT2a$^{551–679}$ crystals were optimized at 4˚C in sitting drop vapour diffusion experiments in 48-well plates, using drops of overall volume 400 nL, mixing protein and mother liquor in a 3:1 protein:mother liquor ratio. *Bs*KKT2 central domain crystals grew from the protein at 26 mg/mL and mother liquor 40% PEG 400, 0.2 mM $(NH_4)_2SO_4$, and 100 mM Tris-HCl pH 8. The 40% PEG400 in the mother liquor served as the cryoprotectant when flash-cooling the crystals by plunging into liquid nitrogen. *Pk*KKT2a$^{551–679}$ crystals grew from the protein at 13 mg/mL and mother liquor 19% MPD, 50 mM Hepes pH 7.5, and 10 mM $MgCl_2$. The crystals were briefly transferred into a cryoprotecting solution of 30% MPD, 50 mM Hepes pH 7.5, 10 mM $MgCl_2$ prior to flash-cooling.

Data collection and structure determination

X-ray diffraction data from a *Bs*KKT2 central domain crystal were collected at the I04 beamline at the Diamond Light Source (Harwell, UK) at the Zinc K-edge wavelength (λ=1.28297 Å). A set of 1441 images were processed in space group I222 using the Xia2 pipeline (Winter, 2010), with DIALS for indexing and integration (Winter et al., 2018) and AIMLESS for scaling (Evans and Murshudov, 2013) to 1.8 Å resolution. Initial 3 Zn atoms were localized by interpreting the anomalous difference Patterson, single anomalous dispersion (SAD) phases were estimated using Crank2 (Skubák and Pannu, 2013), and an initial model was built with BUCCANEER (Cowtan, 2006). The structure was completed by several cycles of alternating model building in Coot (Emsley et al., 2010) and refinement in autoBUSTER (Blanc et al., 2004; Bricogne et al., 2017).

*Pk*KKT2a$^{551–679}$ X-ray diffraction data were collected at the I03 beamline at Diamond Light Source (Harwell, UK) also at the zinc K-edge (λ=1.28272 Å) and processed using the autoPROC pipeline (Vonrhein et al., 2011) using XDS (Kabsch, 2010) for indexing/integration and AIMLESS (Evans and Murshudov, 2013) for scaling, to a resolution of 3.8 Å. Two initial Zn positions were determined by interpreting the anomalous difference Patterson, and single anomalous dispersion (SAD) phases were estimated using Crank2 (Skubák and Pannu, 2013) and SHARP (Vonrhein et al., 2007) in space group $P6_4$. An initial model was manually built in Coot and refined once with

RosettaMR (Terwilliger et al., 2012). The structure was completed by several cycles of alternating model building in Coot (Emsley et al., 2010) and refinement in autoBUSTER (Blanc et al., 2004; Bricogne et al., 2017).

A higher resolution dataset was collected from a *Pk*KKT2a$^{551–679}$ crystal at the I24 beamline at Diamond Light Source (Harwell, UK), at a wavelength of λ=0.9686 Å. Data were processed using Xia2 pipeline (Winter, 2010), DIALS (Winter et al., 2018), and AIMLESS (Evans and Murshudov, 2013) in space group $P6_4$ to a resolution of 2.9 Å. The model obtained from the 3.8 Å dataset was used for further model building and refinement with autoBuster (Bricogne et al., 2017) and Coot (Emsley et al., 2010).

All the images were made with Pymol (Schrödinger LLC, Portland, OR) and CCP4mg (McNicholas et al., 2011). Topology diagrams were generated using TopDraw (Bond, 2003). Protein coordinates have been deposited in the RCSB Protein Data Bank (http://www.rcsb.org/) with accession codes 6TLY (*Bodo saltans* KKT2) and 6TLX (*Perkinsela* KKT2a).

Fluorescent anisotropy DNA-binding assay

All experiments were performed in binding buffer (25 mM Hepes pH 7.5, 50 mM NaCl and 1 mM TCEP) using 1 nM FAM-labelled dsDNA sequences purchased from IDT (Table S4). Prior to the assay, proteins were buffer exchanged into binding buffer using a Zeba spin desalting column (Thermo Fisher), serially diluted at 2:3 v/v ratio and then incubated with DNA for 20 min at room temperature. Fluorescence anisotropy was measured at 25°C using a PHERAstar FS next-generation microplate reader (BMG-Labtech). Each data point is an average of 3 independent experiments. Data were fitted with SigmaPlot using a standard four-parameter logistic equation to calculate K_d.

Acknowledgments

We thank Pietro Roversi and Matt Higgins for advice, crystallography facility manager Edward Lowe, Micron Advanced Bioimaging Unit, and Advanced Proteomics Facility. We also thank Pietro Roversi, Matt Higgins, Danny Huang, and Patryk Ludzia for comments on the manuscript. Midori Ishii was

supported by a long-term fellowship from the TOYOBO Biotechnology Foundation. Bungo Akiyoshi was supported by a Wellcome Trust Senior Research Fellowship (grant no. 210622/Z/18/Z) and the European Molecular Biology Organization Young Investigator Program. The authors declare that no competing interests exist.

Author contributions

Gabriele Marcianò purified recombinant proteins, solved crystal structures, and performed DNA-binding assays. Midori Ishii performed RNAi and rescue experiments. Olga Nerusheva expressed KKT2 and KKT3 truncations and mutants in trypanosomes and performed immunoprecipitation and mass spectrometry. Bungo Akiyoshi performed RNAi experiments and expressed KKT2 and KKT3 truncations and mutants in trypanosomes. Gabriele Marcianò, Midori Ishii, and Bungo Akiyoshi wrote the manuscript.

Author ORCIDs

Gabriele Marcianò, https://orcid.org/0000-0003-2720-652X

Midori Ishii, https://orcid.org/0000-0001-9597-9458

Bungo Akiyoshi, https://orcid.org/0000-0001-6010-394X

Data availability

All data generated during this study are included in the manuscript and supplementary data. Protein coordinates have been deposited in the RCSB Protein Data Bank (http://www.rcsb.org/) with accession codes 6TLY (*Bodo saltans* KKT2) and 6TLX (*Perkinsela* KKT2a).

References

Akiyoshi, B., and Gull, K. (2014). Discovery of unconventional kinetochores in kinetoplastids. Cell *156*, 1247–1258.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. *44*, W344-350.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., et al. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. *38*, D457-462.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., et al. (2005). The genome of the African trypanosome *Trypanosoma brucei*. Science *309*, 416–422.

Black, B. E., and Cleveland, D. W. (2011). Epigenetic centromere propagation and the nature of CENP-A nucleosomes. Cell *144*, 471–479.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M., and Bricogne, G. (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D Biol. Crystallogr. *60*, 2210–2221.

Bond, C. S. (2003). TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics *19*, 311–312.

Brayer, K. J., and Segal, D. J. (2008). Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell Biochem. Biophys. *50*, 111–131.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C., et al. (2017). BUSTER version 2.10.3. Cambridge, United Kingdom: Global Phasing Ltd.

Brun, R., and Schönenberger (1979). Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. Acta Trop. *36*, 289–292.

Brusini, L., D'Archivio, S., McDonald, J., and Wickstead, B. (2019). Ndc80/Nuf2-like protein KKIP1 connects a stable kinetoplastid outer kinetochore complex to the inner kinetochore and responds to metaphase tension. bioRxiv, 764829.

Butenko, A., Opperdoes, F. R., Flegontova, O., Horák, A., Hampl, V., Keeling, P., Gawryluk, R. M. R., Tikhonenkov, D., Flegontov, P., and Lukeš, J. (2020). Evolution of metabolic capabilities and molecular features of diplonemids, kinetoplastids, and euglenids. BMC Biol. *18*, 23.

Carmena, M., Wheelock, M., Funabiki, H., and Earnshaw, W. C. (2012). The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nat. Rev. Mol. Cell Biol. *13*, 789–803.

Cheeseman, I. M., and Desai, A. (2008). Molecular architecture of the kinetochore-microtubule interface. Nat. Rev. Mol. Cell Biol *9*, 33–46.

Colón-González, F., and Kazanietz, M. G. (2006). C1 domains exposed: from diacylglycerol binding to protein-protein interactions. Biochim. Biophys. Acta *1761*, 827–837.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. *62*, 1002–1011.

D'Archivio, S., and Wickstead, B. (2017). Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. J. Cell Biol. *216*, 379–391.

Das, J., and Rahman, G. M. (2014). C1 domains: structure and ligand-binding properties. Chem. Rev. *114*, 12108–12131.

Dathan, N., Zaccaro, L., Esposito, S., Isernia, C., Omichinski, J. G., Riccio, A., Pedone, C., Di Blasio, B., Fattorusso, R., and Pedone, P. V. (2002). The Arabidopsis SUPERMAN protein is able to specifically bind DNA through its single Cys 2 –His 2 zinc finger motif. Nucleic Acids Res *30*, 4945–4951.

Eddy, S. R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Evans, P. R., and Murshudov, G. N. (2013). How good are my data and what is the resolution? Acta Cryst D *69*, 1204–1214.

Findeisen, P., Mühlhausen, S., Dempewolf, S., Hertzog, J., Zietlow, A., Carlomagno, T., and Kollmar, M. (2014). Six subgroups and extensive recent duplications characterize the evolution of the eukaryotic tubulin protein family. Genome Biol Evol *6*, 2274–2288.

French, B. T., and Straight, A. F. (2017). The Power of *Xenopus* Egg Extract for Reconstitution of Centromere and Kinetochore Function. Prog Mol Subcell Biol *56*, 59–84.

Garcia-Silva, M.-R., Sollelis, L., MacPherson, C. R., Stanojcic, S., Kuk, N., Crobu, L., Bringaud, F., Bastien, P., Pagès, M., Scherf, A., et al. (2017). Identification of the centromeres of *Leishmania major*: revealing the hidden pieces. EMBO Reports *18*, 1968–1977.

Hamilton, G. E., and Davis, T. N. (2020). Biochemical evidence for diverse strategies in the inner kinetochore. Open Biol *10*, 200284.

Hara, M., and Fukagawa, T. (2018). Kinetochore assembly and disassembly during mitotic entry and exit. Curr. Opin. Cell Biol. *52*, 73–81.

Henikoff, S., Ahmad, K., and Malik, H. S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098–1102.

Holm, L., and Laakso, L. M. (2016). Dali server update. Nucleic Acids Res. *44*, W351-355.

Ishii, M., and Akiyoshi, B. (2020). Characterization of unconventional kinetochore kinases KKT10 and KKT19 in *Trypanosoma brucei*. J. Cell. Sci. *133*.

Jackson, A. P., Otto, T. D., Aslett, M., Armstrong, S. D., Bringaud, F., Schlacht, A., Hartley, C., Sanders, M., Wastling, J. M., Dacks, J. B., et al. (2016). Kinetoplastid Phylogenomics Reveals the Evolutionary Innovations Associated with the Origins of Parasitism. Curr. Biol. *26*, 161–172.

Jantz, D., and Berg, J. M. (2010). Probing the DNA-binding affinity and specificity of designed zinc finger proteins. Biophys J *98*, 852–860.

Jones, N. G., Thomas, E. B., Brown, E., Dickens, N. J., Hammarton, T. C., and Mottram, J. C. (2014). Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. PLoS Pathog. *10*, e1003886.

Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., et al. (2018). Improvements to the APBS biomolecular solvation software suite. Protein Science *27*, 112–128.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinformatics *20*, 1160–1166.

Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B., et al. (2007). Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. Mol. Biochem. Parasitol. *154*, 103–109.

Kixmoeller, K., Allu, P. K., and Black, B. E. (2020). The centromere comes into focus: from CENP-A nucleosomes to kinetochore connections with the spindle. Open Biol *10*, 200051.

Krishna, S. S., Majumdar, I., and Grishin, N. V. (2003). Structural classification of zinc fingersSURVEY AND SUMMARY. Nucleic Acids Res *31*, 532–550.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. *33*, W299-302.

d'Avila-Levy, C. M., Boucinha, C., Kostygov, A., Santos, H. L. C., Morelli, K. A., Grybchuk-Ieremenko, A., Duval, L., Votýpka, J., Yurchenko, V., Grellier, P., et al. (2015). Exploring the environmental diversity of kinetoplastid flagellates in the high-throughput DNA sequencing era. Mem. Inst. Oswaldo Cruz *110*, 956–965.

Llauró, A., Hayashi, H., Bailey, M. E., Wilson, A., Ludzia, P., Asbury, C. L., and Akiyoshi, B. (2018). The kinetoplastid kinetochore protein KKT4 is an unconventional microtubule tip-coupling protein. J. Cell Biol. *217*, 3886–3900.

London, N., and Biggins, S. (2014). Signalling dynamics in the spindle checkpoint response. Nat. Rev. Mol. Cell Biol. *15*, 736–747.

Lowell, J. E., and Cross, G. A. M. (2004). A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. J. Cell. Sci. *117*, 5937–5947.

Lowell, J. E., Kaiser, F., Janzen, C. J., and Cross, G. A. M. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei*. J. Cell. Sci. *118*, 5721–5730.

Ludzia, P., Lowe, E. D., Marcianò, G., Mohammed, S., Redfield, C., and Akiyoshi, B. (2020). Structural characterisation of KKT4, an unconventional microtubule-binding kinetochore protein. bioRxiv, 2020.10.14.337170.

McKinley, K. L., and Cheeseman, I. M. (2016). The molecular basis for centromere identity and function. Nat. Rev. Mol. Cell Biol. *17*, 16–29.

McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. Acta Cryst D *67*, 386–394.

Musacchio, A., and Desai, A. (2017). A Molecular View of Kinetochore Assembly and Function. Biology (Basel) *6*, 5.

Nerusheva, O. O., and Akiyoshi, B. (2016). Divergent polo box domains underpin the unique kinetoplastid kinetochore. Open Biol *6*, 150206.

Nerusheva, O. O., Ludzia, P., and Akiyoshi, B. (2019). Identification of four unconventional kinetoplastid kinetochore proteins KKT22-25 in *Trypanosoma brucei*. Open Biol *9*, 190236.

Obado, S. O., Bot, C., Nilsson, D., Andersson, B., and Kelly, J. M. (2007). Repetitive DNA is associated with centromeric domains in *Trypanosoma brucei* but not *Trypanosoma cruzi*. Genome Biol. *8*, R37.

Obado, S. O., Taylor, M. C., Wilkinson, S. R., Bromley, E. V., and Kelly, J. M. (2005). Functional mapping of a trypanosome centromere by chromosome fragmentation identifies a 16-kb GC-rich transcriptional "strand-switch" domain as a major feature. Genome Res. *15*, 36–43.

Omichinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M., and Clore, G. M. (1997). The solution structure of a specific GAGA factor–DNA complex reveals a modular binding mode. Nature Structural Biology *4*, 122–132.

Parsons, M., Worthey, E. A., Ward, P. N., and Mottram, J. C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. BMC Genomics *6*, 127.

Poon, S. K., Peacock, L., Gibson, W., Gull, K., and Kelly, S. (2012). A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. Open Biol *2*, 110037.

Siegel, T. N., Hekstra, D. R., and Cross, G. A. M. (2008). Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. Mol. Biochem. Parasitol. *160*, 171–174.

Siegel, T. N., Hekstra, D. R., Kemp, L. E., Figueiredo, L. M., Lowell, J. E., Fenyo, D., Wang, X., Dewell, S., and Cross, G. A. M. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev. *23*, 1063–1076.

Skubák, P., and Pannu, N. S. (2013). Automatic protein structure solution from weak X-ray data. Nat Commun *4*, 2777.

Stankovic, A., and Jansen, L. E. T. (2017). Quantitative Microscopy Reveals Centromeric Chromatin Stability, Size, and Cell Cycle Mechanisms to Maintain Centromere Homeostasis. Prog. Mol. Subcell. Biol. *56*, 139–162.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification *41*, 207–234.

Suzuki, M. (1989). SPKK, a new nucleic acid-binding unit of protein found in histone. EMBO J. *8*, 797–804.

Talbert, P. B., Bayes, J. J., and Henikoff, S. (2009). Evolution of Centromeres and Kinetochores: A Two-Part Fugue. In The Kinetochore:, P. De Wulf and W. C. Earnshaw, eds. (Springer New York), pp. 1–37. Available at: http://www.springerlink.com/content/q810p641t4242254/abstract/.

Tanifuji, G., Cenci, U., Moog, D., Dean, S., Nakayama, T., David, V., Fiala, I., Curtis, B. A., Sibbald, S. J., Onodera, N. T., et al. (2017). Genome sequencing reveals metabolic and cellular interdependence in an amoeba-kinetoplastid symbiosis. Sci Rep *7*, 11688.

Terwilliger, T. C., Dimaio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. V., and Echols, N. (2012). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. J. Struct. Funct. Genomics *13*, 81–90.

UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. *47*, D506–D515.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. *364*, 215–230.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D Biol. Crystallogr. *67*, 293–302.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.

Wickstead, B., Ersfeld, K., and Gull, K. (2002). Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. Mol. Biochem. Parasitol. *125*, 211–216.

Wickstead, B., and Gull, K. (2011). The evolution of the cytoskeleton. J. Cell Biol. *194*, 513–525.

Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. J Appl Cryst *43*, 186–190.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., et al. (2018). DIALS: implementation and evaluation of a new integration package. Acta Cryst D *74*, 85–97.

Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct *29*, 183–212.

Woodward, R., and Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. J. Cell. Sci. *95 (Pt 1)*, 49–57.

Yu, B., Martins, I. R. S., Li, P., Amarasinghe, G. K., Umetani, J., Fernandez-Zapico, M. E., Billadeau, D. D., Machius, M., Tomchick, D. R., and Rosen, M. K. (2010). Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1. Cell *140*, 246–256.

Table 1

Data collection	*Bs*KKT2 central domain
Beamline	Diamond I04
Wavelength (Å)	1.28297
Space group (Z)	I222 (8)
Unit cell (cell edges in Å, cell angles in degrees)	38.96, 53.51, 83.29 90, 90, 90
Resolution range (Å)	45.02 – 1.8 (1.83 – 1.8)
Total No of reflections	37266 (925)
No. of unique reflections	8158 (334)
Completeness (%)	96.96 (77.86)
R_{merge}	0.10 (0.52)
R_{pim}	0.05 (0.33)
$CC_{1/2}$	0.99 (0.77)
$<I/\sigma(I)>$	8.65 (1.08)
Multiplicity	4.57 (2.77)
Anomalous completeness (%)	89.83 (37.37)
Anomalous multiplicity	2.54 (1.93)
Overall *B* factor from Wilson plot (Å^2)	22.4
Refinement	
Resolution (Å)	35.3 – 1.8 (2.0 – 1.8)
No. of reflections working set	7748 (390)
No. of reflections test set	403 (18)
Final R_{work} (%)	19.5 (24.3)
Final R_{free} (%)	22.5 (23.4)
No. of protein atoms	770
No. of Zn atoms	3
No. of water atoms	94
No. of sulfate ions	2
Average B factor (Å^2) protein atoms	30.45
Average B factor (Å^2) Zn atoms	25.31
Average B factor (Å^2) water atoms	44.43
Rmsd bond lengths (Å)	0.008
Rmsd bond angles (°)	0.96
Ramachandran plot	
Most favoured (%)	97.89
Allowed (%)	2.11
Disallowed	0

In parentheses the values relative to the highest resolution shell.

Table 2

Data collection	*Pk*KKT2a$^{551–679}$	
Beamline	Diamond I24	Diamond I03
Wavelength (Å)	0.96861	1.28272
Space group (Z)	$P6_4$ (6)	$P6_4$ (6)
Unit cell (cell edges in Å, cell angles in degrees)	113.84, 113.84, 46.01, 90, 90, 120	114.33, 114.33, 46.32 90, 90, 120
Resolution range (Å)	56.92 – 2.87 (2.92 - 2.87)	99.01 – 3.80 (3.86 - 3.80)
Total No of reflections	153076 (6844)	68026 (3882)
No. of unique reflections	7986 (357)	3535 (189)
Completeness (%)	99.73 (91.77)	100.00 (100.00)
R_{merge}	0.10 (1.04)	0.31 (6.13)
R_{pim}	0.024 (0.240)	0.073 (1.38)
$CC_{1/2}$	0.99 (0.45)	0.99 (0.4)
$<I/\sigma(I)>$	16.45 (2.14)	12 (1.7)
Multiplicity	19.17 (19.17)	19.2 (20.5)
Anomalous completeness (%)	99.6 (91.2)	100 (100)
Anomalous multiplicity	10 (9.9)	10.2 (10.6)
Overall *B* factor from Wilson plot (Å^2)	105.95	136.7
Refinement		
Resolution (Å)	49 – 2.87 (3 – 2.87)	
No. of reflections working set	7408 (417)	
No. of reflections test set	377 (16)	
Final R_{work} (%)	25.2 (36.6)	
Final R_{free} (%)	27.5 (51.2)	
No. of protein atoms	832	
No. of Zn atoms	2	
Average B factor (Å^2) protein atoms	113.4	
Average B factor (Å^2) Zn atoms	100.6	
Rmsd bond lengths (Å)	0.009	
Rmsd bond angles (°)	1.11	
Ramachandran plot		
Most favoured (%)	90.9	
Allowed (%)	9.1	
Disallowed	0	

In parentheses the values relative to the highest resolution shell.

Figure legends

Figure 1. Kinetochore localization of KKT2 and KKT3 is not dependent on various kinetochore proteins

(A) Left: Growth curves of control (gray dashed line) and RNAi-induced cultures (green line) show that these kinetochore proteins are important for proper cell growth (except KKT22) in *T. brucei* procyclic form cells. RNAi was induced with 1 μg/mL doxycycline. Cultures were diluted at day 2. Right: Depletion of indicated kinetochore proteins was confirmed by microscopy. Cells were fixed at either day 1 (RNAi against KKT1, KKT6, KKT7 5′UTR, KKT8 3′UTR, KKT10/19, KKT14, KKT24, and KKIP1) or day 2 (RNAi against KKT22 and KKT23). *T. brucei* has a kinetoplast (K) that contains mitochondrial DNA, and a nucleus (N) that contains nuclear DNA. Kinetoplasts segregate prior to the nuclear division, and the number of kinetoplasts and nuclei can be used as a cell cycle marker (Woodward and Gull, 1990; Siegel et al., 2008). YFP signal was efficiently depleted in > 60% of 2K1N cells (G2 to metaphase) in each case (n > 50, each). Cell lines, BAP672, BAP699, BAP2001, BAP2002, BAP139, BAP2086, BAP1840, BAP1842, BAP1843, BAP770.

(B) Kinetochore localization of KKT2 is not affected by depletion of various kinetochore proteins. RNAi against indicated kinetochore proteins was induced in cells expressing YFP-KKT2. Dot formation was observed in >94% of 2K1N cells in all cases (n > 70, each). Cells were fixed as in (A), except for KKT23 RNAi cells that were fixed at day 4. Cell lines, BAP2004, BAP2005, BAP2006, BAP2007, BAP2008, BAP680, BAP2010, BAP2011, BAP2012, BAP2013.

(C) Kinetochore localization of KKT3 is not affected by depletion of various kinetochore proteins. RNAi against indicated kinetochore proteins was induced in cells expressing KKT3-YFP. Dot formation was observed in >97% of 2K1N cells in all cases (n > 50, each). Cells were fixed as in (B). Cell lines, BAP2014, BAP2015, BAP2016, BAP2017, BAP2018, BAP2085, BAP2020, BAP2021, BAP2022, BAP2023. Scale bars, 5 μm.

Figure 2. KKT2 and KKT3 are important for kinetochore assembly

(A) Growth curve for KKT2 5′UTR RNAi in cells expressing KKT2-YFP. RNAi was induced with 1 μg/mL doxycycline and cultures were diluted at day 2. KKT2-YFP signal was depleted in 46% of 2K1N cells (G2 to metaphase) at day 1 post-induction (n > 100). Cell line, BAP1752.

(B) Quantification of 2K1N cells that had defective kinetochore localization of indicated kinetochore proteins upon induction of KKT2 5′UTR RNAi (n > 100, each). Cells were fixed at day 1 post-induction. In each case, normal kinetochore signal was confirmed in un-induced 2K1N cells (not shown). Cell lines, BAP1743, BAP1753, BAP1749, BAP2075, BAP1750, BAP1746, BAP1751, BAP2080, BAP1747.

(C) Examples of cells expressing indicated kinetochore proteins fused with YFP, showing that KKT3 still forms kinetochore-like dots, while KKT14 fails to localize at kinetochores upon KKT2 depletion.

(D) Growth curve for KKT3 3′UTR RNAi in cells expressing YFP-KKT3. Cultures were diluted at day 2. YFP-KKT3 signal was depleted in 61% of 2K1N cells at day 2 post-induction (n > 100). Cell line, BAP1659.

(E) Quantification of 2K1N cells that had defective kinetochore localization of indicated kinetochore proteins upon induction of KKT3 3′UTR RNAi for 2 days (n > 100, each). In each case, normal kinetochore signal was confirmed in un-induced 2K1N cells (not shown). Cell lines, BAP1755, BAP1764, BAP1761, BAP2076, BAP1762, BAP1758, BAP1763, BAP2081, BAP1759.

(F) Examples of cells expressing indicated kinetochore proteins fused with YFP, showing that KKT2 still forms kinetochore-like dots, while KKT22 and KKT23 failed to localize at kinetochores upon KKT3 depletion.

(G) Growth curve for KKT2/3 double RNAi that targets KKT2 5′UTR and KKT3 3′UTR in cells expressing KKT2-YFP and tdTomato-KKT3. Both KKT2 and KKT3 signals were depleted in 55% of 2K1N cells at day 1 post-induction (n > 100). Cell line, BAP2089.

(H) Quantification of 2K1N cells that had defective kinetochore localization of indicated kinetochore proteins upon induction of KKT2/3 double RNAi for 1 day (n > 100, each). In each case, normal kinetochore signal was confirmed in un-induced 2K1N cells (not shown). Cell lines, BAP2068, BAP2070, BAP2074, BAP2071.

(I) Examples of cells expressing indicated kinetochore proteins fused with fluorescent proteins, showing that KKT1 and KKT4 failed to form normal kinetochore-like dots but instead formed bright blobs upon KKT2/3 depletion. Cells were fixed at day 1 post-induction. Scale bars, 5 µm.

Figure 3. KKT2 has multiple domains that can promote centromere localization in *T. brucei*

(A) Schematic of the *T. brucei* KKT2 protein. Percentages of GFP-positive 2K1N cells (G2 to metaphase) that have kinetochore-like dots were quantified at 1 day post-induction (n > 22, each).

(B) Ectopically expressed *Tb*KKT2 fragments that contain either the central domain (562–677), 672–1030, or the divergent polo boxes (1024–1260) form kinetochore-like dots. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP327, BAP328, BAP381, BAP331, BAP457, BAP519, BAP517. Scale bar, 5 µm.

(C) *Tb*KKT2$^{558–679}$ does not co-purify robustly with other kinetochore proteins. Cell line, BAP382.

(D) *Tb*KKT2$^{672–1030}$ co-purifies with KKIP1 and several other kinetochore proteins. Cell line, BAP519.

(E) *Tb*KKT2 DPB WT, not W1048A, co-purifies with several kinetochore proteins. Cell lines, BAP517, BAP535. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline, and immunoprecipitation was performed using anti-GFP antibodies. See Table S1 for all proteins identified by mass spectrometry.

Figure 4. KKT3 central domain is able to localize at centromeres constitutively in *T. brucei*

(A) Schematic of the *T. brucei* KKT3 protein. Percentages of GFP-positive 2K1N cells (G2 to metaphase) that have kinetochore-like dots were quantified at 1 day post-induction (n > 24, each).

(B) Ectopically expressed *Tb*KKT3 fragments that contain the central domain form kinetochore-like dots. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP291, BAP292, BAP379, BAP296, BAP378, BAP377, BAP418. Scale bar, 5 µm.

(C) *Tb*KKT3 DPB (831–1058) forms kinetochore-like dots during anaphase (88% of 2K2N cells, n = 25). Note that kinetochore proteins typically localize near the leading edge of separating chromosomes during anaphase. Cell line, BAP296. Scale bar, 5 µm.

(D) *Tb*KKT3 DPB co-purifies with several kinetochore proteins. Cell line, BAP520.

(E) *Tb*KKT3$^{594–811}$ does not co-purify robustly with other kinetochore proteins. Cell line, BAP377. See Table S1 for all proteins identified by mass spectrometry.

Figure 5. Crystal structure of *Bodo saltans* KKT2 central domain reveals the presence of two zinc-binding domains

(A) Topology diagram of the *Bs*KKT2 central domain showing the CL domain in green and C2H2-type zinc finger in cyan.

(B) Cartoon representation of the *Bs*KKT2 central domain in two orientations. Zinc ions are shown in grey spheres. The structure is colored as in (A).

(C) Surface representation of the *Bs*KKT2 central domain colored according to sequence conservation using the ConSurf server (Landau et al., 2005; Ashkenazy et al., 2016). Structure orientation as in (B).

(D) Electrostatic surface potential of the *Bs*KKT2 central domain generated by APBS (Jurrus et al., 2018). Structure orientation as in (B).

Figure 6. *Bodo saltans* KKT2 CL is a unique domain

(A) Structure superposition of *Bs*KKT2 CL domain in green and the Vav1 C1 domain in brown (PDB: 3KY9 (Yu et al., 2010)). Zinc ions are shown in grey spheres.

(B) Topology diagram of *Bs*KKT2 CL domain and Vav1 C1 domain.

(C) Close-up view showing a key difference in zinc coordination between *Bs*KKT2 CL domain and Vav1 C1 domain.

Figure 7. Crystal structure of *Perkinsela* KKT2a$^{551–679}$ highlights conservation of CL domain

(A) Cartoon representation of *Pk*KKT2a$^{551–679}$ in two orientations. Zinc ions are shown in grey spheres.

(B) Topology diagram of *Pk*KKT2a$^{551–679}$ structure.

(C) Structure superposition of *Pk*KKT2a$^{551–679}$ and *Bs*KKT2 CL domain, showing that the core of the structure is conserved. Variations between the two structures are due to sequence insertions within CL and the absence of the C2H2 zinc finger at the C-terminus in *Pk*KKT2a$^{551–679}$.

(D) Fluorescence polarization assay for *Pk*KKT2a$^{551–679}$ on 50-bp DNA probes of different GC contents, showing that it binds DNA in a sequence-independent manner. 50 mer CEN probe is part of the centromeric sequence (CIR147) in *T. brucei* and has 36% GC content. Probes, BA1674, BA2218, BA2216.

(E) Fluorescence polarization assay for *Pk*KKT2a$^{551–679\ (C646\ C649A)}$, showing that the mutant has similar DNA-binding affinity. Probe, BA1674.

(F) Fluorescence polarization assay for designed zinc finger (Jantz and Berg, 2010), showing that it has sequence-specific DNA-binding activity. 20 mer CEN DNA has 35% GC content. Probes, BA3083, BA1793, BA1674.

Figure 8. KKT2 CL domain is critical for the centromere localization in *T. brucei*

(A) Multiple sequence alignment of KKT2. Residues that coordinate zinc ions, the conserved aspartic acid residue, and secondary structures of *Bs*KKT2 are shown. Putative SPKK motifs are highlighted in boxes. Numbers below the alignment are for the C2H2 zinc finger's α-helix, where -1 is the residue immediately preceding the helix.

(B) Kinetochore localization of *Tb*KKT2$^{562–677}$ depends on the CL domain, but not on C2H2 zinc finger. Percentages of GFP-positive 2K1N cells that have kinetochore-like dots were quantified at 1 day post-induction (n = 40, each). Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP457, BAP1700, BAP1702, BAP1710, BAP1712, BAP1715, BAP1717, BAP1649, BAP1719, BAP1837.

(C) *Tb*KKT2 C576A and D622A mutants localize at kinetochores but fail to support normal cell growth. One allele of *Tb*KKT2 was mutated and tagged with an N-terminal YFP tag, and the other allele was depleted using RNAi-mediated knockdown by targeting the 5′UTR of the *Tb*KKT2 transcript. Top: cells were diluted every two days and cell growth was monitored for 8 days upon induction of RNAi. Controls are uninduced cell cultures. Similar results were obtained for at least three clones of *Tb*KKT2 mutants. Bottom: Example of cells expressing the *Tb*KKT2 mutants prior to RNAi induction, showing that they localize at kinetochores (n > 150, each) (also see Figure S7A).

Maximum intensity projections are shown. RNAi was induced with 1 μg/mL doxycycline. Cell lines, BAP1789, BAP1779, BAP1786. Scale bars, 5 μm.

Figure 9. Kinetochore localization of KKT3 depends on the central domain in *T. brucei*

(A) Multiple sequence alignment of KKT3. Residues that are expected to coordinate zinc ions as well as the conserved aspartic acid residue are shown.

(B) Percentage of GFP-positive cells that have kinetochore-like dots were quantified at 1 day post-induction (n > 22, each). Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP291, BAP359, BAP360, BAP446, BAP447, BAP1721, BAP1722, BAP362, BAP341.

(C) *Tb*KKT3 C668A/C671A and D692A mutants do not localize at kinetochores and fail to support normal cell growth, while *Tb*KKT3 C707A/C710A mutant is functional. One allele of *Tb*KKT3 was mutated and tagged with a C-terminal YFP tag, and the other allele was depleted using RNAi-mediated knockdown by targeting the 3′UTR of the *Tb*KKT3 transcript. Top: cells were diluted every two days and cell growth was monitored for 8 days upon induction of RNAi. Similar results were obtained for at least three clones of *Tb*KKT3 mutants. Controls are uninduced cell cultures. Bottom: Example of cells expressing the *Tb*KKT3 mutants prior to RNAi induction, showing that *Tb*KKT3$^{C668A\ C671A}$ and *Tb*KKT3^{D692A} do not localize at kinetochores while *Tb*KKT3$^{C707A\ C710A}$ localizes normally (n > 90, each) (also see Figure S7B). Maximum intensity projections are shown. RNAi was induced with 1 μg/mL doxycycline. Cell lines, BAP1791, BAP1793, BAP1783. Scale bars, 5 μm.

Supplemental Tables (Excel files)

Table S1: List of proteins identified in the immunoprecipitates of YFP-tagged KKT2$^{558–679}$, KKT2$^{672–1030}$, KKT2$^{1024–1260}$, KKT2$^{1024–1260\ (W1048A)}$, KKT3$^{594–811}$, and KKT3$^{846–1058}$ by mass spectrometry

Table S2: List of DALI search hits for the *Bs*KKT2 CL domain structure

Table S3: List of DALI search hits for the *Bs*KKT2 C2H2 zinc finger structure

Table S4: List of trypanosome cell lines, plasmids, primers, synthetic DNA, and DNA probes used in this study

Supplementary Figure legends

Figure S1. KKT2 fragments localize at kinetochores from S phase to anaphase, while KKT3 central domain localizes constitutively

(A–C) Ectopically expressed *Tb*KKT2 central domain (562–677), *Tb*KKT2$^{672–1030}$, and *Tb*KKT2 DPB (1024–1260) localize at kinetochores from S phase until anaphase.

(D) Ectopically expressed *Tb*KKT3 central domain (594–728) forms kinetochore-like dots throughout the cell cycle. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. tdTomato-KKT1 was used as a kinetochore marker. Cell lines, BAP1998, BAP2000, BAP1999, BAP1997. Scale bars, 5 μm.

Figure S2. Purification of *Bodo saltans* KKT2, *Perkinsela* KKT2a, and designed zinc finger proteins

Size exclusion chromatography of (A) *Bodo saltans* KKT2 central domain, (B) *Pk*KKT2a$^{551–679}$, and (C) designed zinc finger with respective SDS-PAGE gels showing pooled fractions. HiPrep Superdex 75 16/60 column was used. Note that two separate gels are shown for (A). M stands for protein marker.

Figure S3. Multiple sequence alignment of KKT2 and KKT3 homologs in kinetoplastids

Residue numbers in *Tb*KKT2 for those cysteines and histidines that coordinate zinc ions in the *Bs*KKT2 structure are listed at the top of the alignment to highlight the conservation of these residues among kinetoplastids. Note that cysteine is used in KKT2 and KKT2-like proteins (C576 in *Tb*KKT2), while histidine is present in KKT3 in slightly different position (highlighted in pink box). Position of the strictly conserved aspartic acid residue is also shown. SPKK motifs (black box) and a putative AT-hook motif in *Perkinsela* KKT2b (red box) are also highlighted.

Figure S4. Full-length KKT2 mutants localize at kinetochores

Full-length KKT2 proteins with indicated mutations were ectopically expressed with 10 ng/mL doxycycline. Cell lines, BAP327, BAP365, BAP366, BAP367, BAP368, BAP369. Scale bar, 5 µm.

Figure S5. *Bodo saltan*s KKT2 CL domain surface sequence conservation and electrostatic potential reveal the presence of a conserved acidic patch centering on residue D626

(A) Surface sequence conservation of *Bs*KKT2 CL domain using the ConSurf server (Landau et al., 2005; Ashkenazy et al., 2016). The C2H2 zinc finger structure is shown as cartoon representation.

(B) Electrostatic surface potential of the *Bs*KKT2 CL domain generated by APBS (Jurrus et al., 2018) reveals the presence of a conserved acidic surface. The location of the residue D626 is marked by a black circle.

Figure S6. RNAi of KKT2 and KKT3 causes growth defects in *T. brucei*

Growth curves of (A) KKT2 5′UTR RNAi, (B) KKT2 5′UTR RNAi with YFP-KKT2 (resistant to the RNAi), (C) KKT3 3′UTR RNAi, and (D) KKT3 3′UTR RNAi with KKT3-YFP (resistant to the

RNAi). 1 μg/mL doxycycline was added to induce RNAi. Controls are uninduced cell cultures. Similar results were obtained from at least two independent experiments. Cell lines, BAP1554, BAP1681, BAP1555, BAP1682.

Figure S7. Kinetochore-like dots of KKT2 and KKT3 mutants co-localize with KKT1

(A) Indicated KKT2 mutants expressed at endogenous locus co-localize with a kinetochore marker, tdTomato-KKT1. Cell lines, BAP2036, BAP2033, BAP2035.

(B) $KKT3^{C707A\ C710A}$ co-localizes with tdTomato-KKT1, while $KKT3^{C668A\ C671A}$ and $KKT3^{D692A}$ mutants do not. Cell lines, BAP2037, BAP2038, BAP2034.

Figure 1

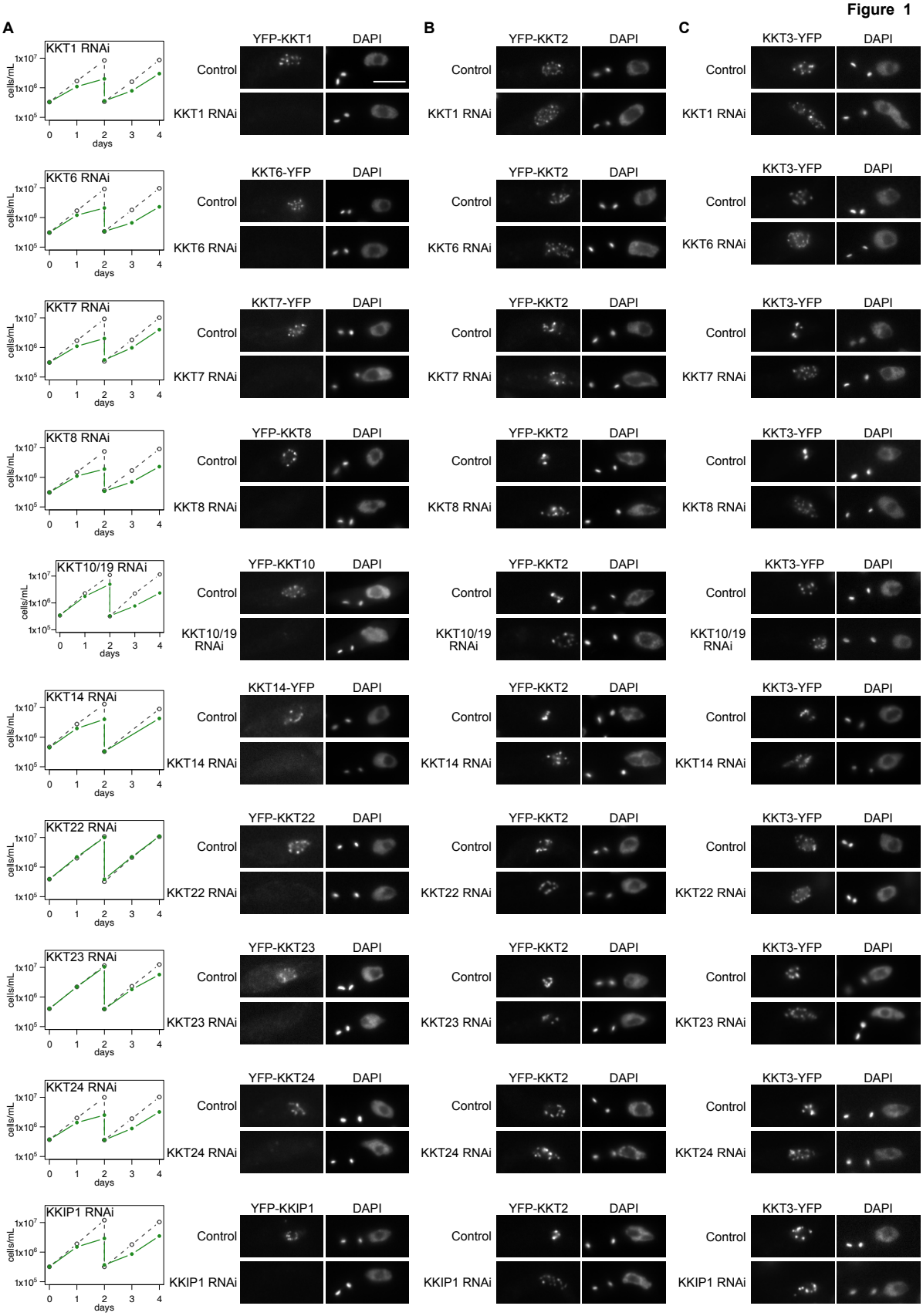

A
B
C
KKT1 RNAi
KKT6 RNAi
KKT7 RNAi
KKT8 RNAi
KKT10/19 RNAi
KKT14 RNAi
KKT22 RNAi
KKT23 RNAi
KKT24 RNAi
KKIP1 RNAi
cells/mL
days
YFP-KKT1
KKT6-YFP
YFP-KKT7
YFP-KKT8
YFP-KKT10
KKT14-YFP
YFP-KKT22
YFP-KKT23
YFP-KKT24
YFP-KKIP1
YFP-KKT2
KKT3-YFP
DAPI
Control

Figure 2

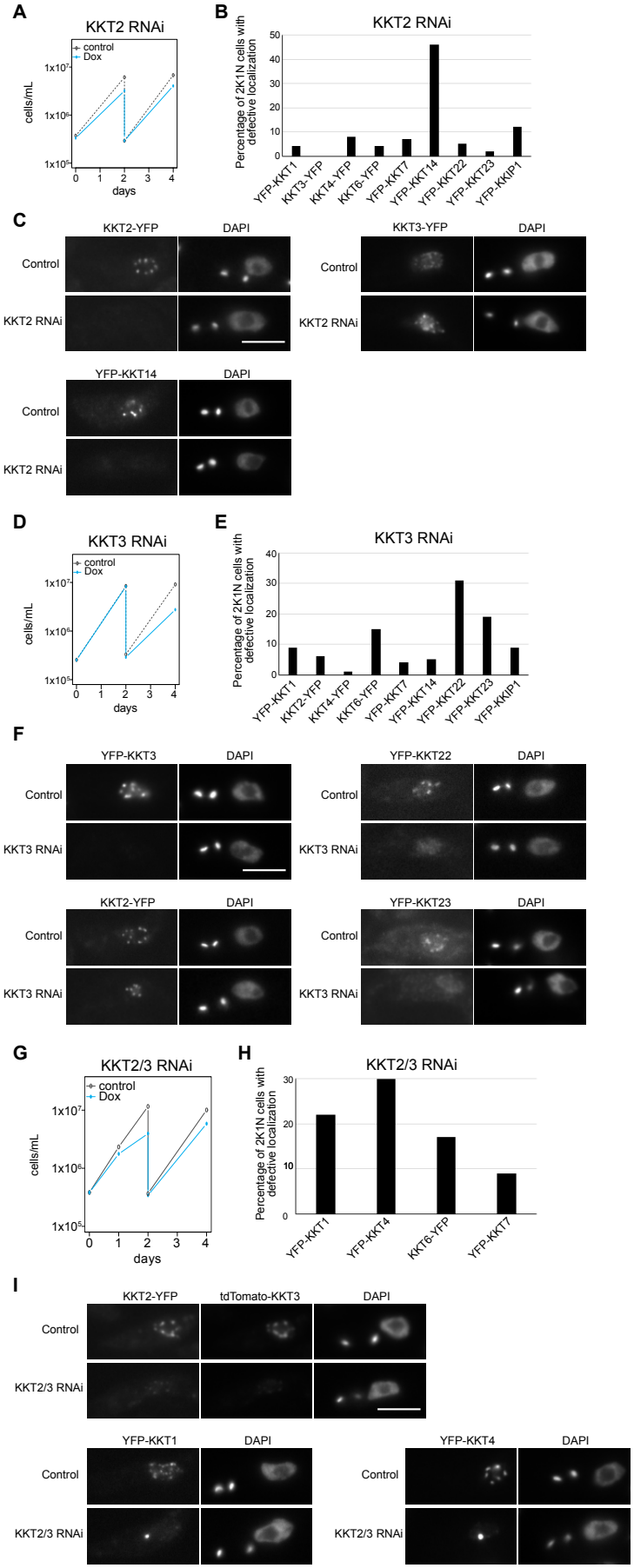
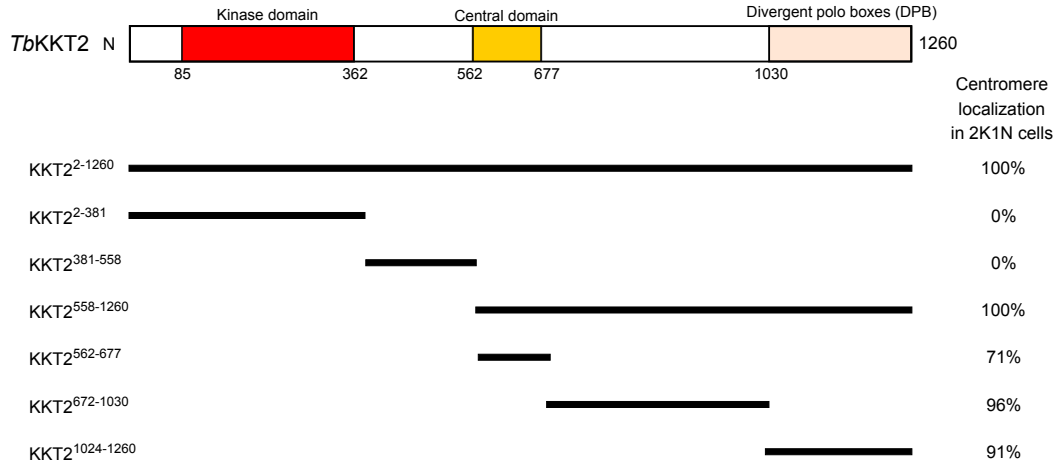
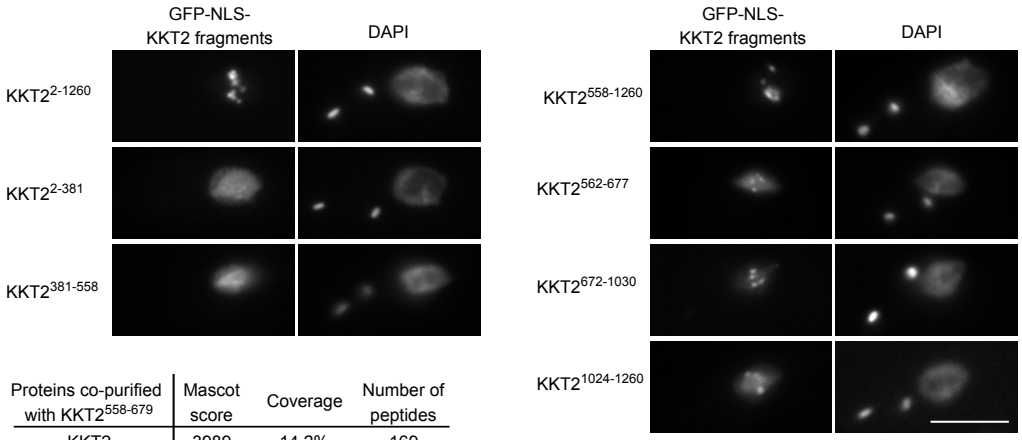

A
KKT2 RNAi
control
Dox
cells/mL
days
B
KKT2 RNAi
Percentage of 2K1N cells with defective localization
YFP-KKT1
KKT3-YFP
KKT4-YFP
KKT6-YFP
YFP-KKT7
YFP-KKT14
YFP-KKT22
YFP-KKT23
YFP-KKIP1
C
KKT2-YFP
DAPI
KKT3-YFP
YFP-KKT14
Control
KKT2 RNAi
D
KKT3 RNAi
E
KKT3 RNAi
KKT2-YFP
F
YFP-KKT3
YFP-KKT22
YFP-KKT23
KKT3 RNAi
G
KKT2/3 RNAi
H
KKT2/3 RNAi
YFP-KKT1
YFP-KKT4
KKT6-YFP
YFP-KKT7
I
tdTomato-KKT3
KKT2/3 RNAi

Figure 3

A



B



C

Proteins co-purified with KKT2 558-679	Mascot score	Coverage	Number of peptides
KKT2	3989	14.2%	169
KKT3	129	12.9%	6
KKT5	45	25.0%	3

D

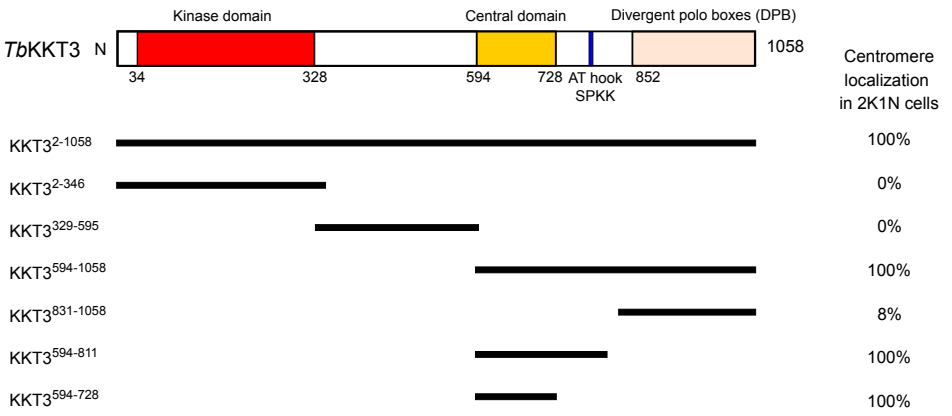
Proteins co-purified with KKT2 672-1030	Mascot score	Coverage	Number of peptides
KKT2	6263	23.7%	343
KKIP1	2574	46.4%	126
KKT3	368	14.3%	18
KKT1	199	6.1%	10
KKT7	144	10.4%	8

E

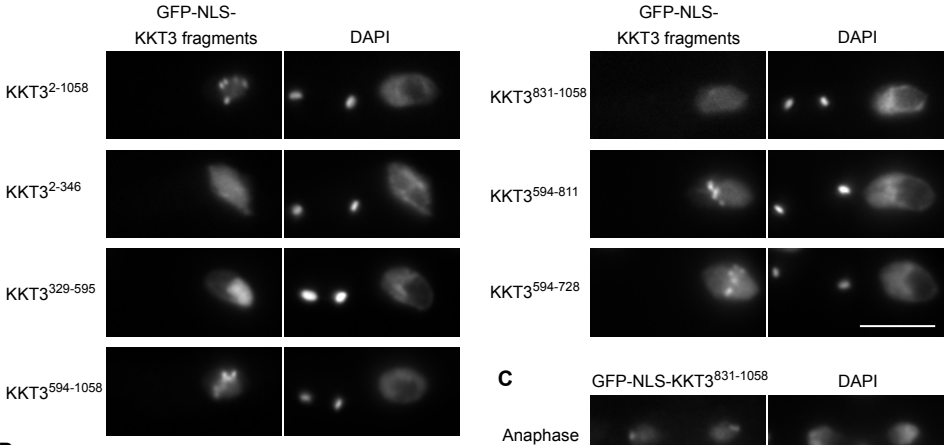
Proteins co-purified with KKT2 1024-1260	WT			W1048A
	Mascot score	Coverage	Number of peptides	Number of peptides
KKT2	2341	22.1%	106	116
KKT1	458	9.8%	22	0
KKT6	105	23.4%	3	0
KKT7	96	7.9%	4	0
KKT8	80	17.5%	7	0

Figure 4

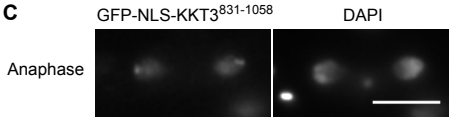
A



B



C



D

Proteins co-purified with KKT3$^{846-1058}$	Mascot score	Coverage	Number of peptides
KKT3	1923	13.5%	123
KKT2	548	20.6%	17
KKT8	368	20.3%	9
KKT5	313	39.0%	11
KKT7	311	20.5%	9
KKT1	197	12.1%	9
KKT19	147	37.6%	6
KKT10	85	30.3%	5

E

Proteins co-purified with KKT3$^{594-811}$	Mascot score	Coverage	Number of peptides
KKT3	2899	20.9%	152
KKT7	174	19.6%	7
KKT1	114	13.6%	5
KKT2	96	17.6%	5

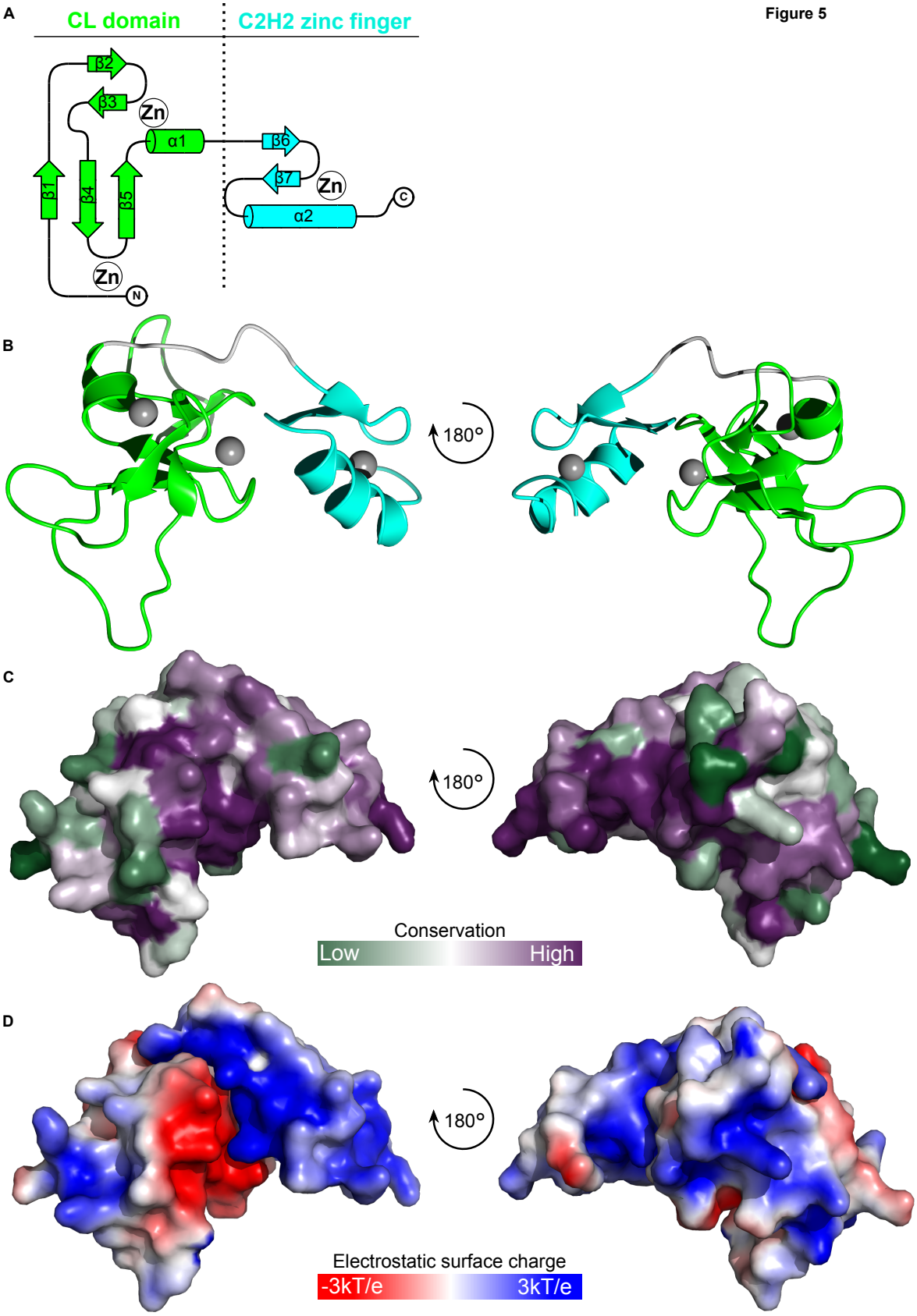

A
CL domain
C2H2 zinc finger
β2
β3
Zn
α1
β6
β7
Zn
α2
C
β1
β4
β5
Zn
N
B
180°
C
180°
Conservation
Low
High
D
180°
Electrostatic surface charge
-3kT/e
3kT/e

Figure 5

Figure 6

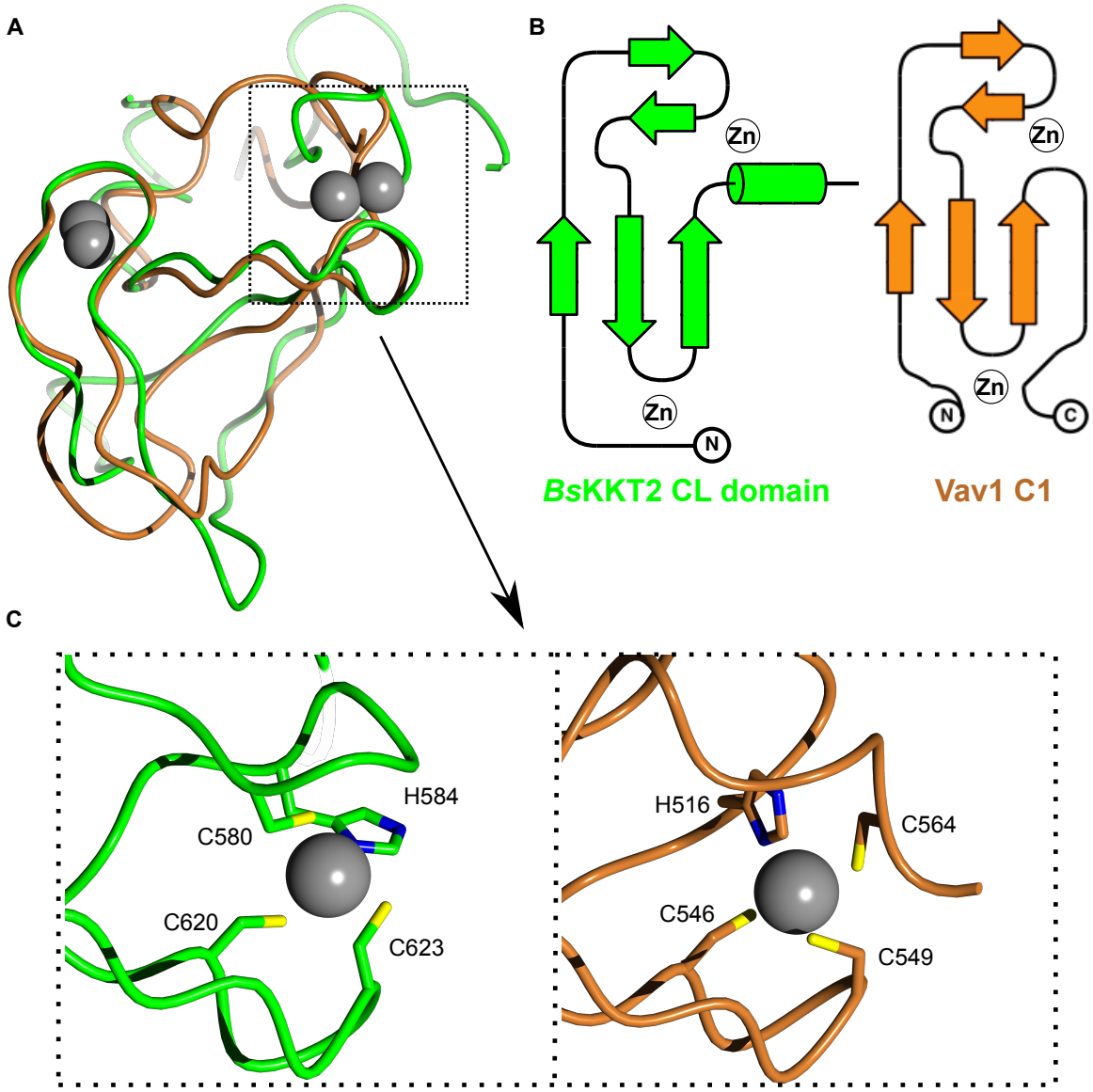

A
B
Zn
Zn
N
BsKKT2 CL domain
Zn
Zn
N
C
Vav1 C1
C
H584
C580
C620
C623
H516
C564
C546
C549

Figure 7

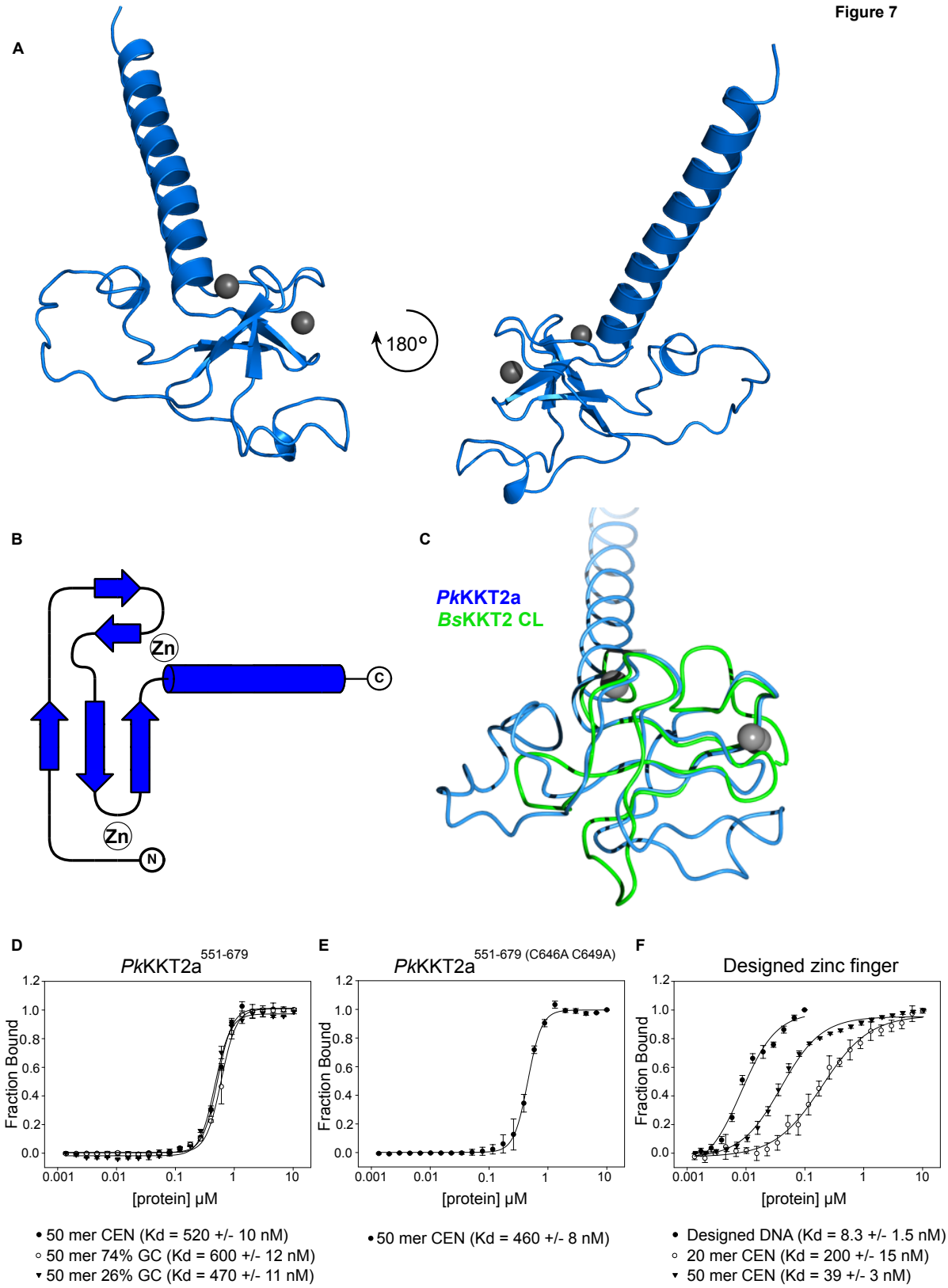

A
180°
B
Zn
C
Zn
N
C
PkKKT2a
BsKKT2 CL
D
PkKKT2a 551-679
Fraction Bound
[protein] µM
• 50 mer CEN (Kd = 520 +/- 10 nM)
◦ 50 mer 74% GC (Kd = 600 +/- 12 nM)
▾ 50 mer 26% GC (Kd = 470 +/- 11 nM)
E
PkKKT2a 551-679 (C646A C649A)
Fraction Bound
[protein] µM
• 50 mer CEN (Kd = 460 +/- 8 nM)
F
Designed zinc finger
Fraction Bound
[protein] µM
• Designed DNA (Kd = 8.3 +/- 1.5 nM)
◦ 20 mer CEN (Kd = 200 +/- 15 nM)
▾ 50 mer CEN (Kd = 39 +/- 3 nM)

Figure 8

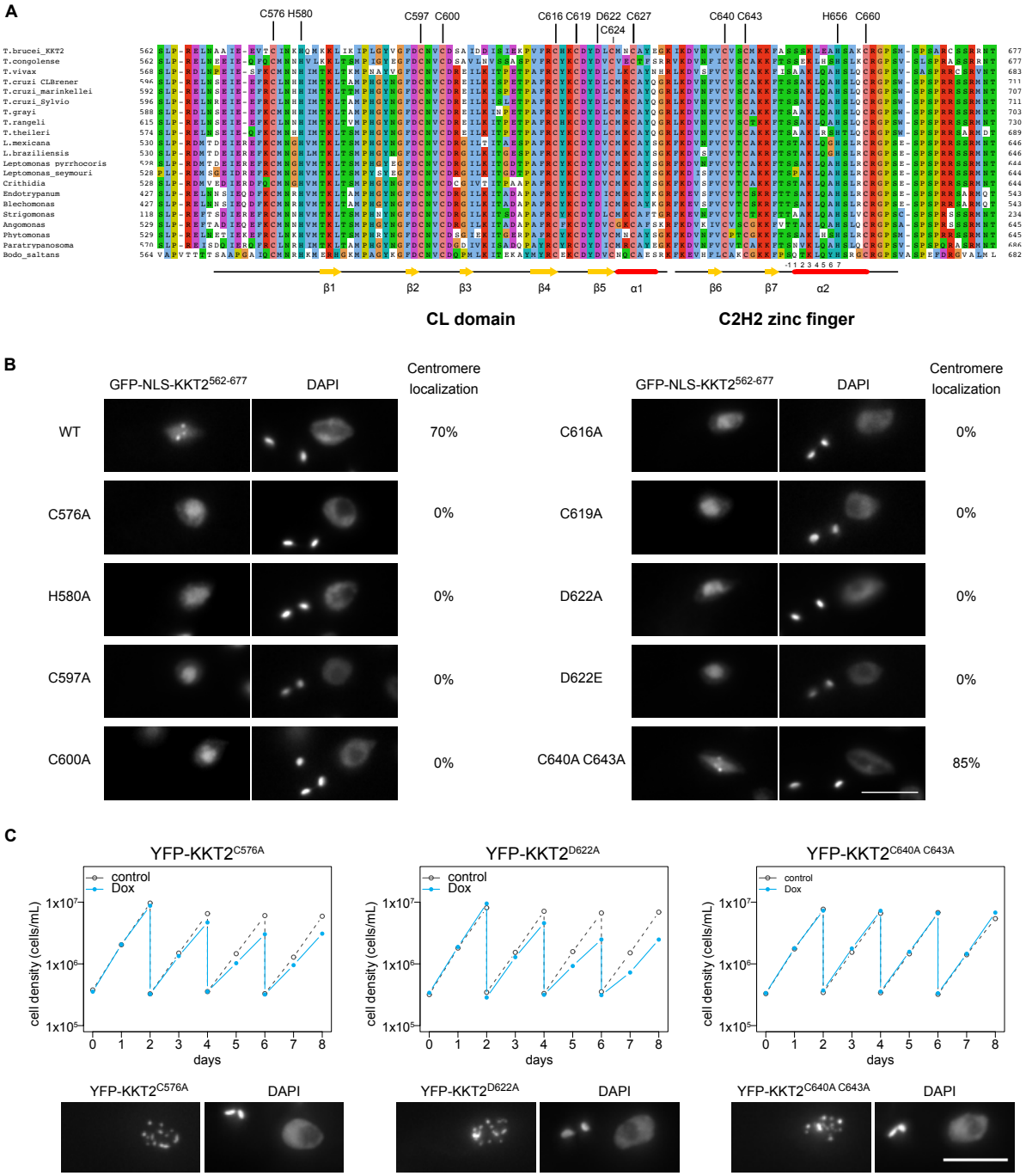
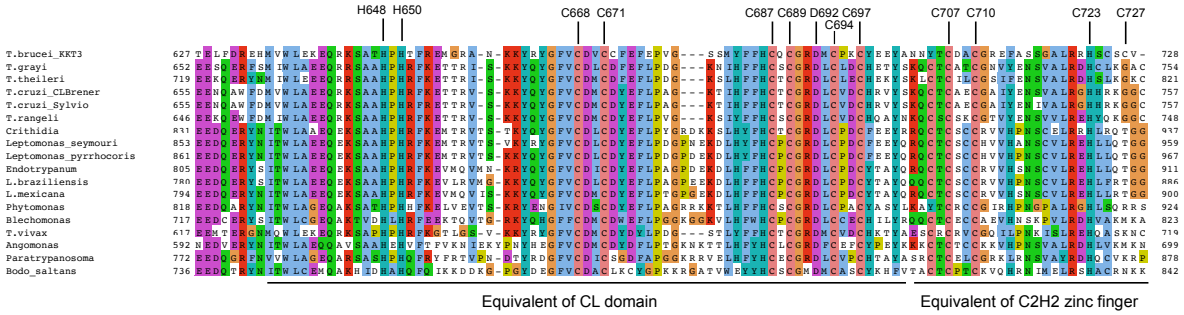
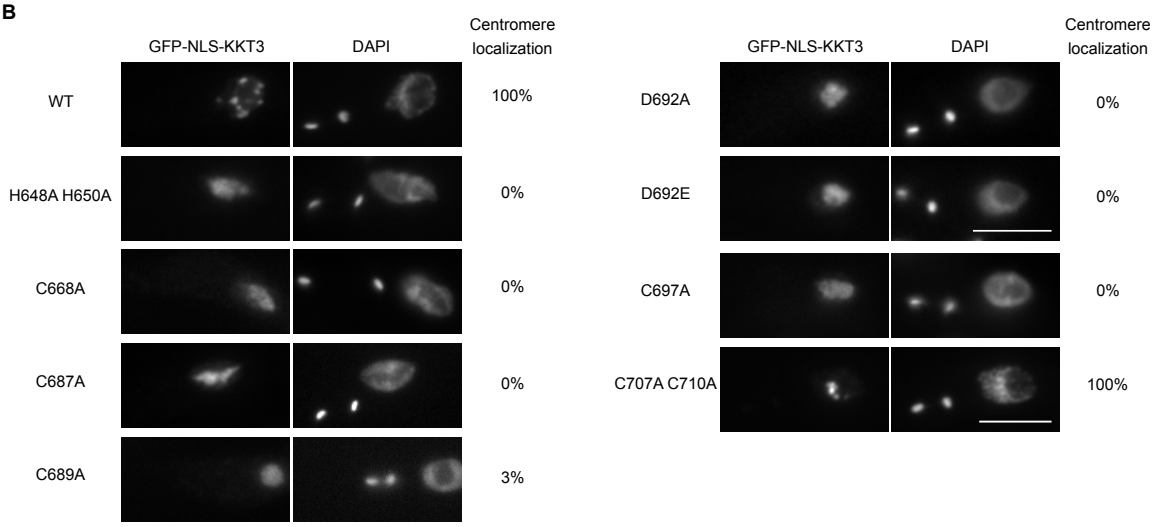

A
C576 H580
C597 C600
C616 C619 D622 C627
C624
C640 C643
H656 C660
T.brucei_KKT2
T.congolense
T.vivax
T.cruzi CLBrener
T.cruzi_marinkellei
T.cruzi_Sylvio
T.grayi
T.rangeli
T.theileri
L.mexicana
L.braziliensis
Leptomonas pyrrhocoris
Leptomonas_seymouri
Crithidia
Endotrypanum
Blechomonas
Strigomonas
Angomonas
Phytomonas
Paratrypanosoma
Bodo_saltans
β1
β2
β3
β4
β5
α1
β6
β7
α2
CL domain
C2H2 zinc finger
B
GFP-NLS-KKT2562-677
DAPI
Centromere localization
WT
70%
C576A
0%
H580A
0%
C597A
0%
C600A
0%
C616A
0%
C619A
0%
D622A
0%
D622E
0%
C640A C643A
85%
C
YFP-KKT2C576A
YFP-KKT2D622A
YFP-KKT2C640A C643A
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
days
0 1 2 3 4 5 6 7 8
DAPI

Figure 9

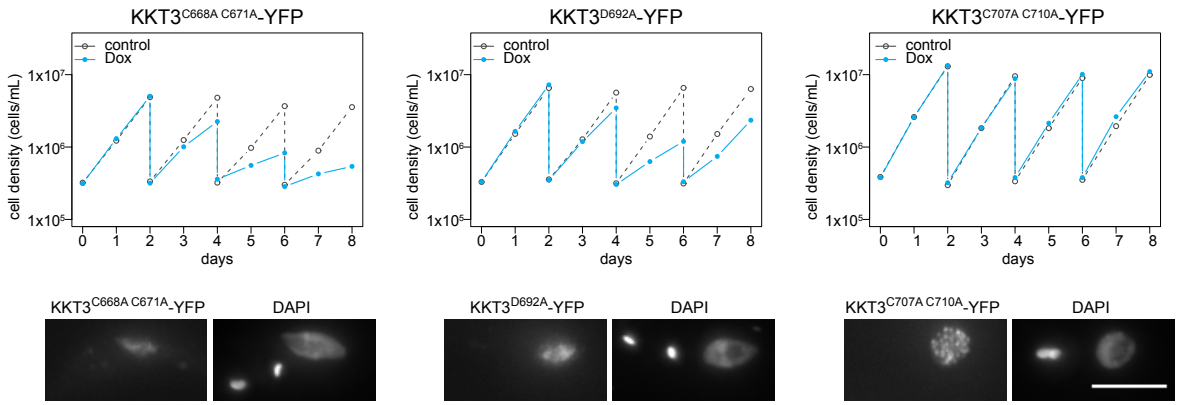
A


H648 H650
C668 C671
C687 C689 D692 C697
C694
C707 C710
C723 C727
T.brucei_KKT3
T.grayi
T.theileri
T.cruzi_CLBrener
T.cruzi_Sylvio
T.rangeli
Crithidia
Leptomonas_seymouri
Leptomonas_pyrrhocoris
Endotrypanum
L.braziliensis
L.mexicana
Phytomonas
Blechomonas
T.vivax
Angomonas
Paratrypanosoma
Bodo_saltans
Equivalent of CL domain
Equivalent of C2H2 zinc finger

B


GFP-NLS-KKT3
DAPI
Centromere localization
WT
100%
H648A H650A
0%
C668A
0%
C687A
0%
C689A
3%
D692A
0%
D692E
0%
C697A
0%
C707A C710A
100%

C


KKT3C668A C671A-YFP
KKT3D692A-YFP
KKT3C707A C710A-YFP
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
days
DAPI

Figures S1

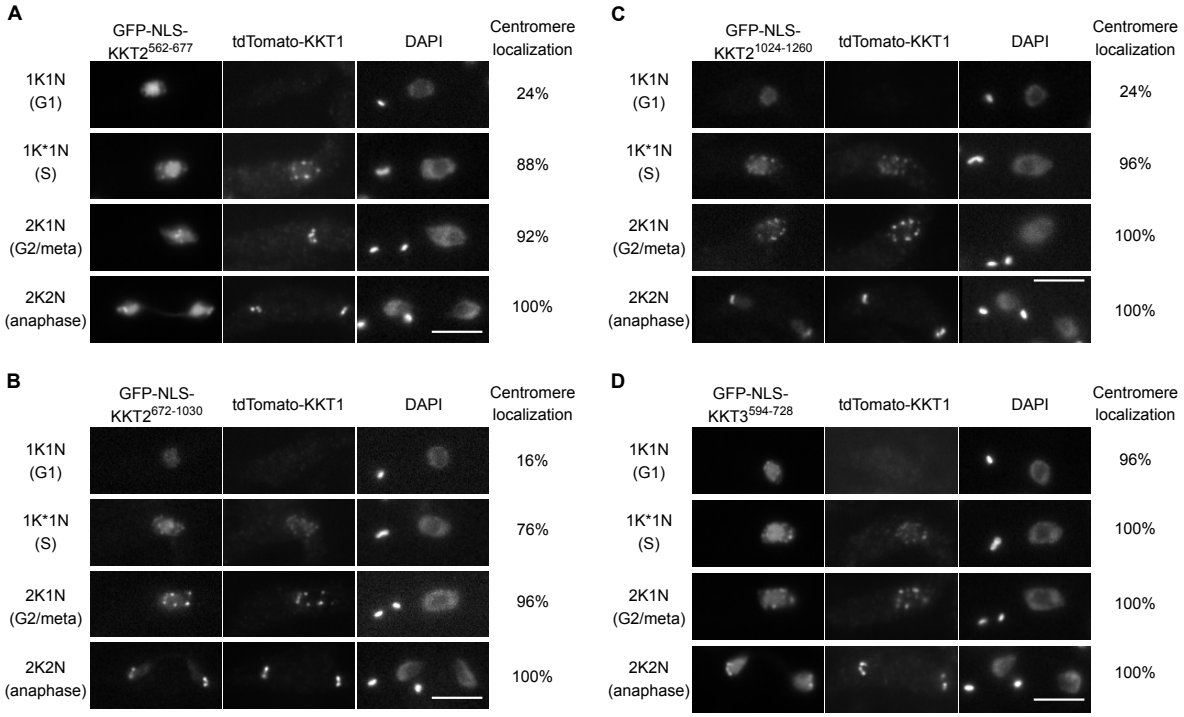

A
GFP-NLS-KKT2562-677
tdTomato-KKT1
DAPI
Centromere localization
1K1N (G1)
24%
1K*1N (S)
88%
2K1N (G2/meta)
92%
2K2N (anaphase)
100%
B
GFP-NLS-KKT2672-1030
tdTomato-KKT1
DAPI
Centromere localization
1K1N (G1)
16%
1K*1N (S)
76%
2K1N (G2/meta)
96%
2K2N (anaphase)
100%
C
GFP-NLS-KKT21024-1260
tdTomato-KKT1
DAPI
Centromere localization
1K1N (G1)
24%
1K*1N (S)
96%
2K1N (G2/meta)
100%
2K2N (anaphase)
100%
D
GFP-NLS-KKT3594-728
tdTomato-KKT1
DAPI
Centromere localization
1K1N (G1)
96%
1K*1N (S)
100%
2K1N (G2/meta)
100%
2K2N (anaphase)
100%

Figure S2

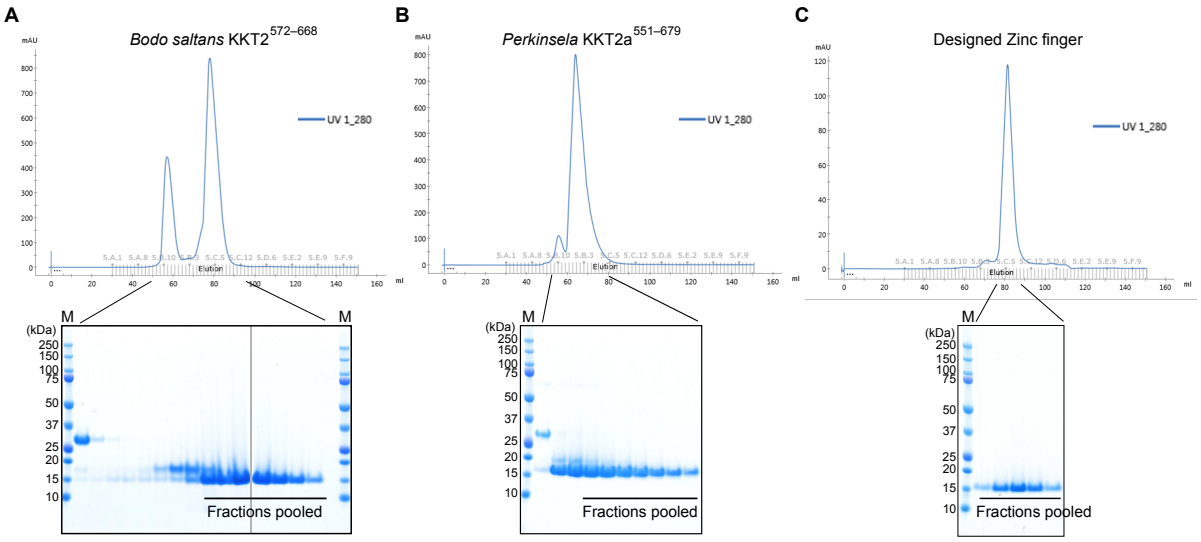

A
Bodo saltans KKT2572–668
UV 1_280
Elution
M
(kDa)
250
150
100
75
50
37
25
20
15
10
Fractions pooled
B
Perkinsela KKT2a551–679
UV 1_280
Elution
M
(kDa)
Fractions pooled
C
Designed Zinc finger
UV 1_280
Elution
M
(kDa)
Fractions pooled

Figure S3

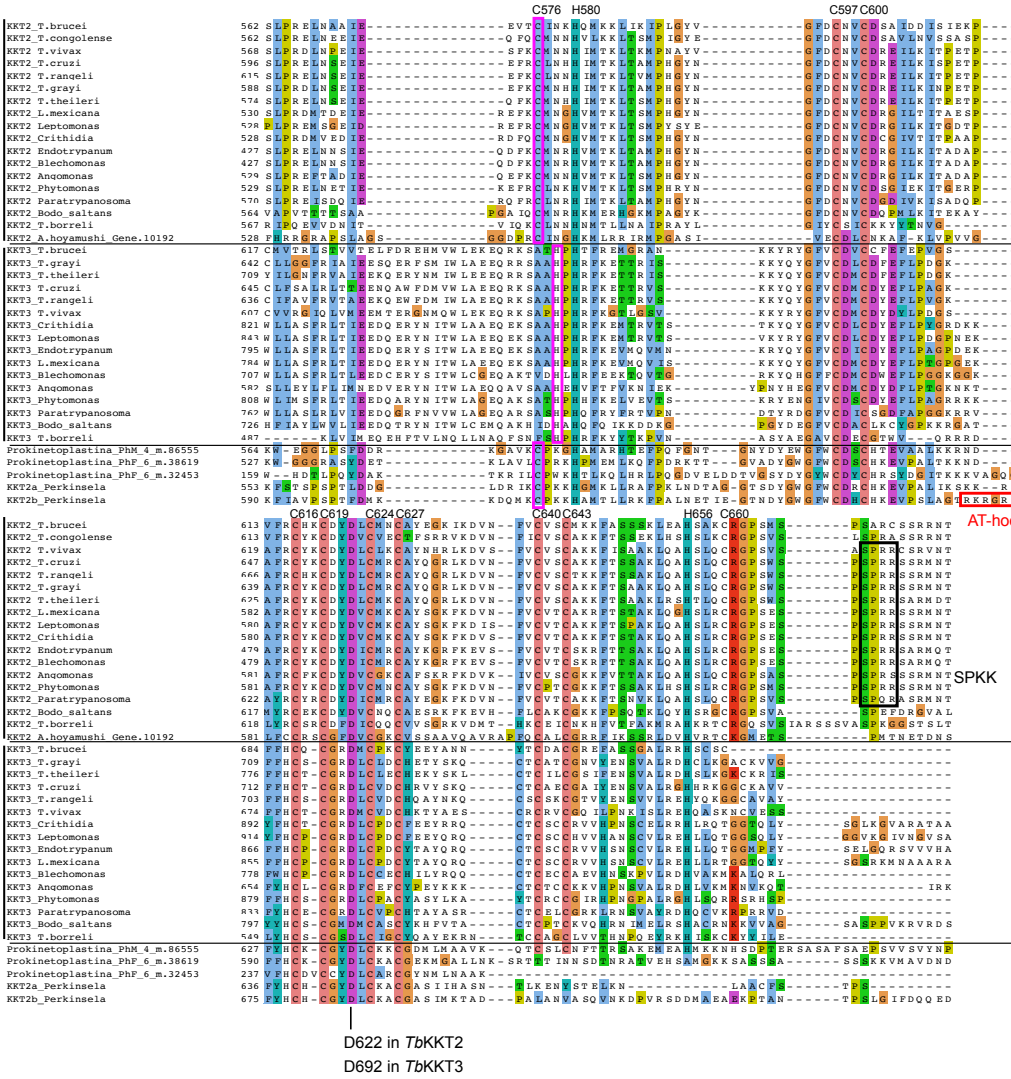

C576 H580
C597 C600
C616 C619 C624 C627
C640 C643
H656 C660
AT-hook
SPKK
D622 in TbKKT2
D692 in TbKKT3

Figure S4

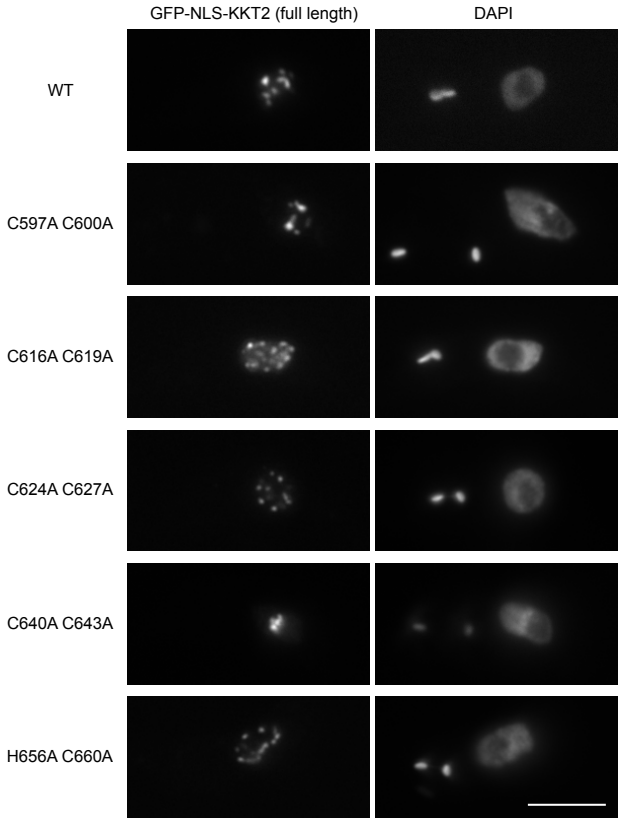

GFP-NLS-KKT2 (full length)
DAPI
WT
C597A C600A
C616A C619A
C624A C627A
C640A C643A
H656A C660A

Figure S5

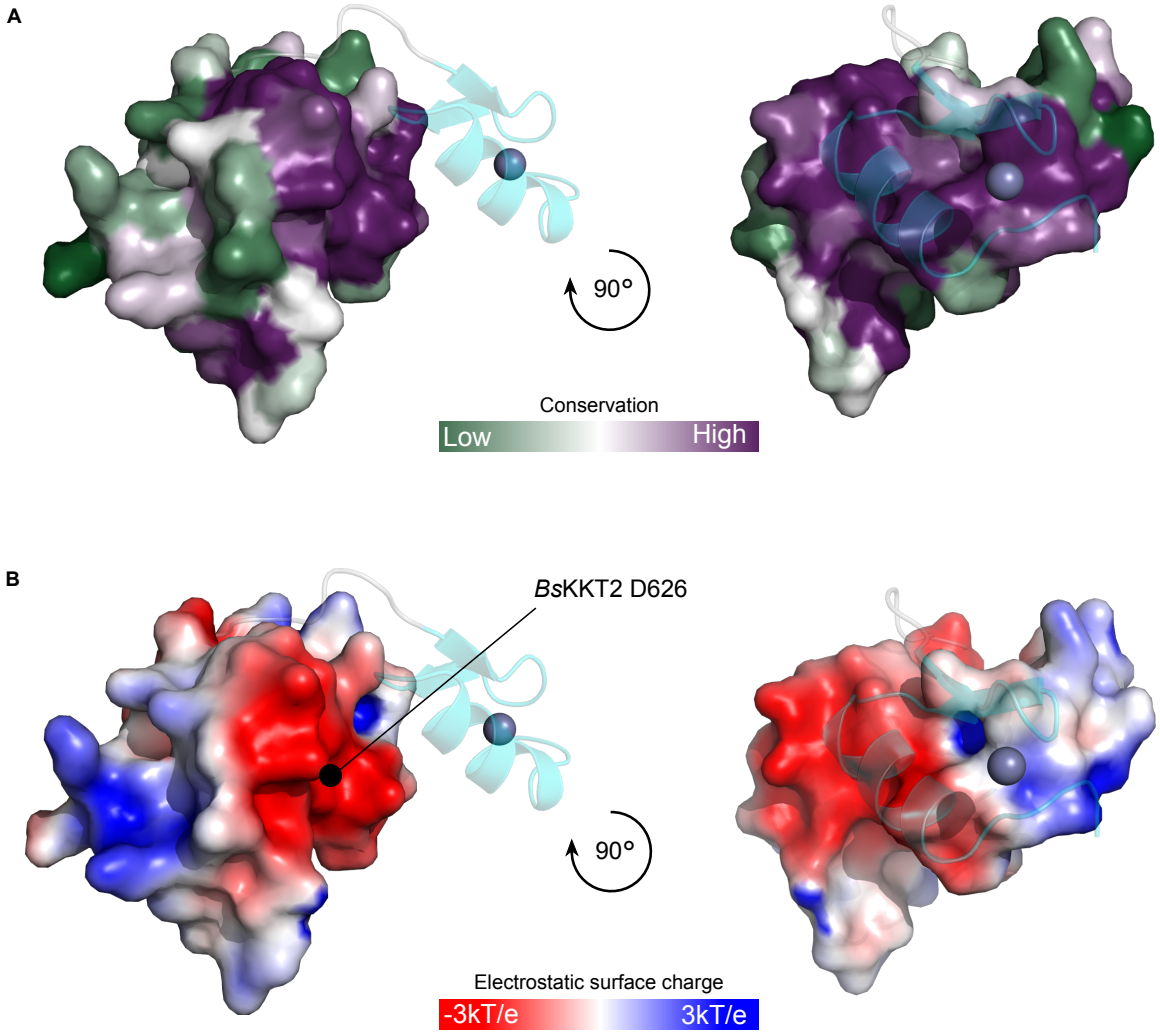

A
90°
Conservation
Low
High
B
*Bs*KKT2 D626
90°
Electrostatic surface charge
-3kT/e
3kT/e

Figure S6

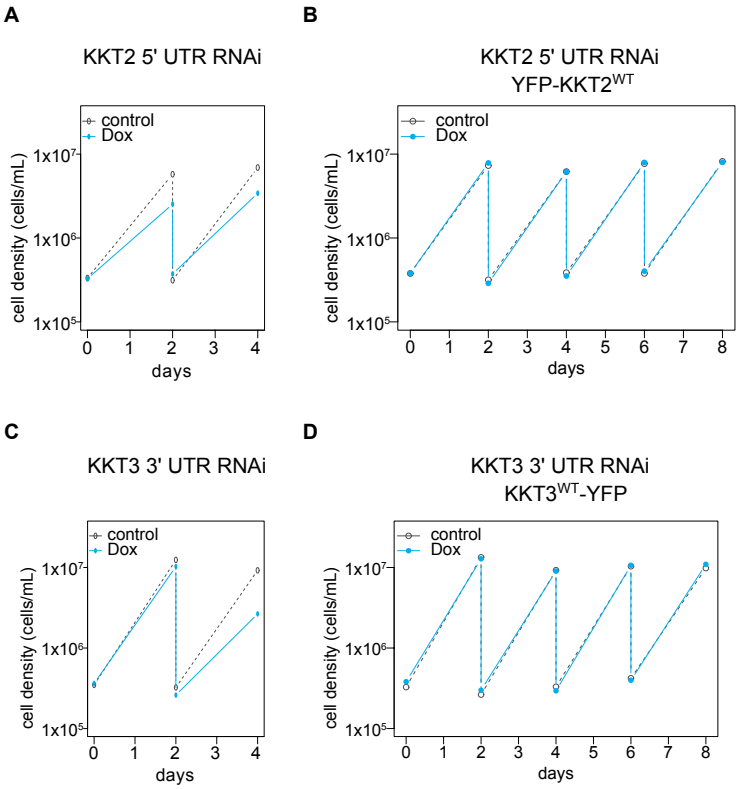
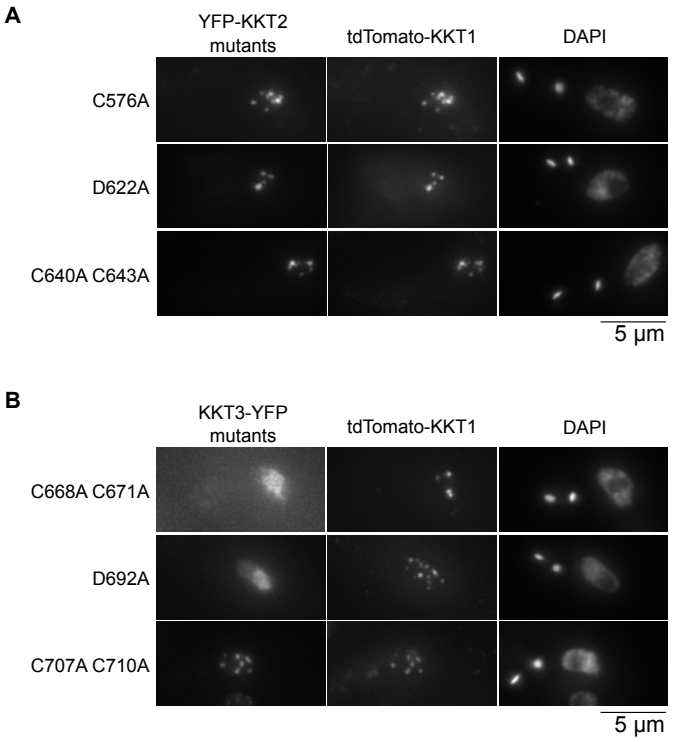

A
KKT2 5' UTR RNAi
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
0 1 2 3 4
days
B
KKT2 5' UTR RNAi
YFP-KKT2^WT
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
0 1 2 3 4 5 6 7 8
days
C
KKT3 3' UTR RNAi
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
0 1 2 3 4
days
D
KKT3 3' UTR RNAi
KKT3^WT-YFP
control
Dox
cell density (cells/mL)
1x10^7
1x10^6
1x10^5
0 1 2 3 4 5 6 7 8
days

Figure S7


A
YFP-KKT2 mutants
tdTomato-KKT1
DAPI
C576A
D622A
C640A C643A
5 µm
B
KKT3-YFP mutants
tdTomato-KKT1
DAPI
C668A C671A
D692A
C707A C710A
5 µm